Task demands determine comparison strategy in whole probe change detection.

Rob Udale

University of Bristol

Simon Farrell

University of Western Australia

Chris Kent

University of Bristol

Word count: 14,897.

Author note.

This research was funded by a "studentship in teaching and research" from the Faculty of Science at the University of Bristol. The second author was supported by an Australian Future Fellowship (FT130100149).

We would like to thank Jeremy Ward, Jack Moss, and Niall Taylor for their assistance with data collection.

Correspondence concerning this article should be addressed to Rob Udale, School of Experimental Psychology, University of Bristol, United Kingdom.

Contact: ru13940@bristol.ac.uk

Abstract

Detecting a change in our visual world requires a process that compares the external environment (test display) with the contents of memory (study display). We addressed the question of whether people strategically adapt the comparison process in response to different decision loads. Study displays of three colored items were presented, followed by 'whole-display' probes containing three colored shapes. Participants were asked to decide whether any probed items contained a new feature. In Experiments 1-4, irrelevant changes to the probed item's locations or feature bindings influenced memory performance, suggesting that participants employed a comparison process that relied on spatial locations. This finding occurred irrespective of whether participants were asked to decide about the whole display, or only a single cued item within the display. In Experiment 5, when the base-rate of changes in the non-probed items increased (increasing the incentive to use the cue effectively), participants were not influenced by irrelevant changes in location or feature bindings. In addition, we observed individual differences in the use of spatial cues. These results suggest that participants can flexibly switch between spatial and non-spatial comparison strategies, depending on interactions between individual differences and task demand factors. These findings have implications for models of visual working memory that assume that the comparison between study and test obligatorily relies on accessing visual features via their binding to location.

Public significance statement.

Detecting changes in the environment relies on a process which compares the external environment with the contents of memory. It is thought that this comparison process

automatically relies on the spatial locations of the items which are being compared. However, we have found that, although a spatial comparison process appears to be the default comparison process, our participants were also able to conduct a non-spatial comparison process. These results suggest that we can flexibly switch between different comparison strategies in order to decide whether a change has occurred in the environment. This may have future practical implications for how we can train individuals to make better use of their working memory resources, for example, by training more effective comparison and decision-making strategies.

Key words: Visual working memory, Change detection, Feature binding, Location binding, Relational encoding, Comparison process.

Many every-day decisions rely, in part, on making a comparison between the contents of memory and the external environment. The manner in which these comparisons are made will depend on the task at hand. For example, in order to choose the best apple at the grocery store, the features of the current best apple need to be temporarily stored in memory, and compared to the features of multiple apples sampled from the stall (Vidal, Gauchou, Tallon-Baurdy, & O'Reagan, 2005). In other cases, such as to re-establish object correspondence between saccades, only the currently fixated item is compared with the memory representation of the previously fixated item (Currie, McConkie, Carolson-Radavanksy, & Irwin, 2000). Contemporary models of visual working memory (VWM) have primarily focused on describing limits in VWM capacity (e.g., Luck & Vogel, 1997; Wilken & Ma, 2004; Vandenberg, Awh & Ma, 2014), and on the format of memory representations, such as whether items are represented as fully integrated objects (Vogel, Woodman & Luck, 2001), or whether features are represented independently (Wheeler & Treisman, 2002). Consequently, there has been less focus on the way in which stored representations are used to guide decision-making. This article addresses strategies by which VWM contents are compared with the external environment to detect changes in the environment.

Many existing conceptualizations of VWM (e.g., Jiang, Olson & Chun, 2000; Hyun, Woodman, Vogel, Hollingworth, & Luck, 2009; Treisman & Zhang, 2006) assume a comparison process that employs *exhaustive in-place matching*. Exhaustive in-place matching is a process with the following assumptions: i) during the encoding of a study display, task-relevant visual features are automatically bound to their spatial locations, irrespective of the task-relevancy of location information (Bodoroglu & Shah, 2009; Golomb, Kupitz, & Thiemann, 2014; Kondo & Saiki, 2012; Oliviers & Schreij, 2014; Treisman & Zhang, 2006); ii) when a test display is presented, each item represented in VWM is automatically compared to the item presented in the corresponding location in the external environment (Jiang et al. 2000); iii) representations are not compared to items at other locations in the scene (Treisman & Zhang, 2006); iv) this comparison process occurs exhaustively, in other words, it occurs for every item represented in the test display (Hyun et al. 2009); and v) the evidence from all of these comparisons is integrated in order to form a decision, such as whether or not a change has occurred in the display (Palmer, 1990; Shaw, 1980; Wilken & Ma, 2004). Central to exhaustive in-place matching is the idea that locations are automatically or obligatorily encoded (Bodoroglu & Shah, 2009; Golomb et al. 2014; Kondo & Saiki, 2012; Oliviers & Schreij, 2014; Treisman & Zhang, 2006). Many papers have provided evidence supporting the assumption that visual features are effortlessly, if not automatically (Bodoroglu & Shah, 2009; Kondo & Saiki, 2012; Treisman & Zhang, 2006), or obligatorily (Hayes, Nadel, & Ryan, 2007; Papenmeier & Huff, 2014; Silvis & Shapiro, 2014), bound to their relative (to other items in the display) spatial locations (Bodoroglu & Shah, 2009; Hollingworth, 2006, 2007, 2009; Jiang et al. 2000; Mou, Xiao, & McNamara, 2008).

Evidence for exhaustive in-place matching comes from the change detection task. In this task, participants must decide if a change has occurred in the features of items between a study display and a subsequent test display. The displays are typically separated by a brief (e.g. 900 ms) maintenance interval consisting of a blank screen. In one often-used version of the task, participants are asked to remember the visual features of the displayed items, but to disregard their locations. Typically, studies have found a reduction in accuracy when the locations in the test display do not match the locations in the study display (Chen, 2009; Guérard, Morey, Lagacé, & Tremblay, 2013; Hollingworth, 2006, 2007; Hollingworth & Rasmussen, 2010; Jiang et al. 2000; Pertzov & Husain, 2014; Treisman & Zhang, 2006, although see Woodman, Vogel & Luck, 2009 for an exception), despite these changes being irrelevant to the instructions of the task. According to an exhaustive in-place matching account, this performance drop occurs because people are able to use exhaustive in-place matching when the locations match. However, when the locations have changed, correspondence cannot be re-established between the objects in the two displays, so people must rely on a less efficient comparison process. For example, because correspondence cannot be established for objects between study and test displays, they might match a probed object to multiple items in memory. Furthermore, when the task instructions are reversed, so that participants must remember where items appeared, but disregard visual features such as color and shape, there is no detrimental effect of changing the task-irrelevant features. This suggests that people are not using the visual features to guide the comparison process for locations in the same way that spatial location is used for the comparison of visual features. Supporting evidence for the use of exhaustive in-place matching also comes from cued recall tasks, in which participants are presented a series of color-orientation feature conjunctions, and asked to recall one feature (e.g., color) when cued with another (e.g., orientation). Participants often erroneously recall the features of uncued items which appeared in the same location as the target during the study display, suggesting that participants automatically retrieve the features associated by their spatial location (Pertzov & Husain, 2014).

Despite the evidence outlined above, there is good reason to believe that the patterns of data supporting a special role for location may instead result from the manner in which participants strategically approach these tasks. First, Woodman et al. (2012) found no effect of changing item locations in a change detection task unless the target item was spatially cued. The difference in effects between cue conditions (present or absent) may therefore be due to the presence of a spatial cue, which encouraged a spatial comparison strategy, whereas a non-spatial comparison may have been employed when the spatial cue was absent. Additionally, Allen, Castellà, Ueno, Hitch and Baddeley (2015) have shown that task-irrelevant spatial information is automatically encoded to memory representations, but only when the task cues for locations, suggesting that the encoding of location is subject to strategic influences. Third, using functional magnetic resonance imaging (fMRI), Vincente-Grabovetsky, Carlin & Cusack (2012) found that when participants knew that a probed item would appear in its original location, there was stronger retinotopic encoding---as indexed by an increased BOLD response in areas v1 and v2 of the visual cortex---suggesting that the locations were being maintained during the task. In contrast, when probed items were presented in new locations on every trial, so that participants could not predict the locations of the probed items, retinotopic coding was reduced during the maintenance period, suggesting that participants did not actively try and maintain spatial information. Finally, Bodoroglu and Shah (2009) split their sample as to whether a participant's false alarm rate (responding 'Change' to an unchanged item) was above or below the median of the sample. For the low false alarm rate group, changing probed object locations substantially reduced performance relative to when probed items were in their original location. However, for the high false alarm rate group, changing the locations of the test display items had no effect on performance, relative to presenting them in their original locations. One interpretation of this finding is that some participants were strategically utilizing location in their comparison process. This strategy provides the benefit of reducing false alarm rates, because it produces fewer erroneous comparisons when the test items are presented in their original locations, at the cost of reduced sensitivity when the probed items have changed their locations. This could be a viable trade-off because the relative spatial locations were kept the same on two-thirds of the trials.

Further evidence that exhaustive in-place matching may not always be employed comes from studies using the single probe design, whereby memory is probed for only a single item rather than for all items in the study display (Gilchrist & Cowan, 2014). The classic finding that response times (RT) linearly increase as a function of set size (Letter recognition: Sternberg, 1966, 1969; Colour change detection: Gilchrist & Cowan, 2014; Hyun et al. 2009) suggests that participants search memory representations in an effortful, if not serial, manner. In other words, the probed item may be compared to multiple items (at different locations) in memory. Gilchrist and Cowan (2014) showed that the set size effect on RT was the same irrespective of whether the probed item was presented in its original location or in a new location. If participants consistently made use of location information in the single probe condition, the search would always be terminated after the first comparison in memory, because it would always match the probed item. As a result, when the probed items were presented in their original locations, the slope of the linear relationship between set size and RT should be flatter than the slope for probe items presented at new locations. The finding that search times across set sizes were the same in both conditions suggests that participants exhaustively searched memory in both conditions, comparing the probed item to item representations at multiple locations, inconsistent with in-place matching.

The inconsistencies in the literature described above raise the question of whether exhaustive in-place matching is indeed an automatic process, or whether it is adaptively tailored to the nature of the task. Our suggestion is that the disruptive effect of changing locations in the change detection task can be explained by participants changing their response strategies in response to task demands, rather than a necessarily automatic comparison process. In the case where all the studied items were probed (whole-display probe) and all items are in their original locations, the participant can employ exhaustive in-place matching. The reason why they might choose this strategy is because each of the probed items only needs to be compared to a single item representation in memory (Figure 1A). In contrast, if participants employed some other comparison process, which was not guided by item locations, they might need to compare each probed item with multiple items in memory, and integrate the accumulated evidence from each comparison to produce a decision (Figure 1B). This strategy is potentially more effortful, and therefore it is reasonable that participants will take advantage of any cue, such as location, which can be used to reduce the number of unnecessary comparisons required to make their decision. Furthermore, this strategic interpretation also provides an account of why performance is unaffected by location changes in single item probes (Gilchrist & Cowan, 2014; Treisman & Zhang, 2006). First, in the case of single probes, comparing the probe and study displays is already less effortful than in whole-display probes. For example, imagine three items are studied, followed by a whole-display probe. If location is not utilized, the participants could potentially make up to nine comparisons (if each probe item is compared to all the studied items). However, if three items are studied, but only a single probe is presented, only a maximum of three comparisons need to be made in order to exhaustively search memory. Therefore, less effort is required in order to compare a single probe to multiple memory representations bound to different locations (Figure 1C) than there is with a whole-display probe. In the case of single probes, less effort is saved by limiting the search of memory to the probed items location (Figure 1D). In the case of single probes, match comparisons are more diagnostic than in full probes, because they only need to check whether the probe matches any of the memory representations. It is plausible, therefore, that when single probes are presented, the optimal strategy---in terms of reducing effort required to make the decision---is to search multiple items represented in memory, irrespective of whether they share locations with the probed item. It seems relatively clear that the detrimental effect of changing locations for whole-display probes, but not in single probes, can be accounted for by strategic shifts in matching processes adopted by participants. If this is the case, then it should be possible to encourage use of different strategies, and thus modulate the disruptive effects of changing location, by varying the potential effectiveness of those strategies.

In a series of experiments, we compared the disruptive effects of location changes on whole-display probes against probes that were otherwise identical, except that only a single cued item was tested. It was expected that presenting a whole-display probe containing a cue which indicated which item was the target should facilitate a non-spatial search, as is thought to be the case when single probes are presented (Gilchrist & Cowan, 2014; Sternberg, 1966), rather than in-place matching. The results across the experiments suggest that in most cases (Experiments 1, 2, and 4), participants employ exhaustive in-place matching, even when given the chance to utilize other comparison strategies. Experiment 3 provides some evidence that individual differences influence whether in-place matching is exhaustive (occurs for all probed items) or partial (occurs for a single, or sub-set of probed items). Finally, in Experiment 5, we further discouraged the use of exhaustive in-place matching by increasing the probability of a to-be-ignored feature change occurring (an irrelevant change in a non-cued test item). In Experiment 5 we found evidence that participants utilized a non-spatial comparison strategy, suggesting that they can strategically switch between comparison strategies, in response to task demands. The stimulus materials, experiment scripts, data, and analysis scripts can be found on the open science framework (osf.io/e7g26).

Experiment 1

In Experiment 1, we presented whole-display probes in a change detection task. On half of the trials, we indicated which item was the target: the item that would change if a change was going to occur, and therefore the only item participants needed to base their decision on. Thus, the decision load for this “Cue present” condition is similar to a single item probe in that the decision only needs to be made about a single test item, rather than the entire display. The critical difference between the probe display in this experiment and a single probe display was that the spatial configuration, composed of the other studied items, was also present on the screen during the retrieval and decision stage. It has already been shown that participants employ a non-spatial comparison process when a single probe is presented (Treisman & Zhang, 2006), but do they also employ this process in this single probe analogue?

To assay which type of comparison process was being used, we based our design on that of Treisman and Zhang (2006). Participants were shown displays containing colored shapes, and asked only to make decisions based on the presence or absence of color or shape changes, but not on locations, or the combinations (bindings) between colors and shapes. When probed, the items were presented in either their original location or in new locations. In addition, the combination of color and shape was either kept the same as in the study display, or swapped between items. Treisman and Zhang (2006) found lower accuracy in the new location condition, relative to the old location condition, when the bindings were intact. However, when the bindings were switched, Treisman and Zhang (2006) found higher accuracy in the new location condition compared to the old location condition. In addition, they found lower overall accuracy in the switched-binding condition than in the intact binding condition. One explanation for the location-binding interaction was that participants were utilizing exhaustive in-place matching: Each probe item was compared to the studied item at the same location, but not to items that were presented at other locations. Therefore, when the bindings had switched places, it appeared to the participants as though each location contained a ‘new’ feature. Furthermore, this interaction did not occur when single probes were presented, possibly because for this probe type, they were not using a space-based comparison process such as in-place matching. Because of the modulation of the interaction by probe type (a three-way interaction between probe type, binding, and location) it was a suitable method for measuring the presence or absence of in-place matching. In all experiments, the presence or absence of the three-way interaction is therefore diagnostic of whether participants were using location in their decision making.

There were two possible outcomes for this experiment. The first was that global in-place matching automatically occurs whenever the studied configuration is presented at test. In this case, we expect participants to use in-place matching irrespective of whether the cue was present or absent, and so the pattern of performance should be equivalent in both the Cue present and Cue absent conditions. Alternatively, if participants can strategically change their response strategies, we expect participants to use in-place matching in the Cue absent condition, consistent with previous findings (Treisman & Zhang, 2006), but to employ a non-spatial comparison process in the Cue present condition (similar to the single probe task; Treisman & Zhang, 2006).

Method

Participants. Because our design contained a replication of Treisman and Zhang (2006, Experiment 1), and we wished to ensure we replicated their basic effect, we determined the sample size using Simonshon's (2015) 'small telescopes' heuristic - that replication studies should multiply the original sample size by 2.5. Therefore, we recruited thirty-two naïve participants (aged 18-32, 26 were female) for each experiment. We conducted a power analysis for the two-way Location x Binding interaction from one of our previous studies (Udale, Farrell & Kent, 2017), which used a similar design and sample size. Based on the effect size from this previous study [$F(1, 28) = 11.47$, $\eta p^2 = 0.65$], we should expect to achieve power of 0.9 for the two-way Location x Binding interaction, with a sample size of 28 participants. Before taking part, we asked participants about their visual acuity. All participants reported normal, or corrected-to-normal, vision. Ethical approval was granted by the University of Bristol, Faculty of Science Research Ethics Committee.

Materials. The presentation of the stimuli was controlled using MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), using a 17" TFT monitor (Resolution: 1,280 x 1,024) with a refresh rate of 60 Hz. Participants' responses were recorded using a standard USB keyboard. The memoranda were presented on a uniform medium grey background (RGB: 128, 128, 128). Each of the study and probe displays consisted of three items, consisting of conjunctions of six possible colors and shapes, without repetitions. The colors were red (RGB: 255, 0, 0), blue (RGB: 0, 0, 255), yellow (RGB: 255, 255, 0), green (RGB: 0, 0, 255), brown (RGB: 150, 75, 0) and violet (RGB: 238, 130, 238). The shapes were a circle, a square, an equilateral triangle, a heart, a star, and a cross. Each shape subtended approximately 1.4° x 1.4° visual angle, at a viewing distance of approximately one meter. Each item could appear in one of nine possible locations. The nine locations formed a 3 x 3 grid, each location consisting of 60 x 60 pixels (1.8° x 1.8° visual angle), with 36 pixels (1.08° visual angle) of empty space between each location. The total grid size was 252 x 252 pixels (7.6° x 7.6° visual angle). The shapes, colors, and locations were randomly chosen at the start of each trial without replacement. Finally, a cue was presented on some trials, consisting of the outline of a square (RGB: 255, 255, 255) centered on the target item, subtending 2.3° x 2.3° degrees of visual angle.

Design and procedure. The experimental design was a fully crossed 2 (Feature: Match vs. Change) x 2 (Location: Old Vs. New) x 2 (Binding: Intact vs. Switched) x 2 (Cue: Present vs. Absent) factorial within subjects design. Participants took part in a practice block of sixteen trials, followed by sixteen experimental blocks.

On each trial, a small white cross appeared at the centre of the screen for 1,000 ms, followed by a study display for 150 ms, which consisted of three color-shape items. The study display was followed by a blank maintenance screen, with a centrally presented fixation cross, for 900 ms. After the maintenance screen, a probe display was presented, containing three color-shape items. The probe display remained on the screen until a response was given by the participant. Participants were asked to decide if the probe display contained a new feature that was not present in the study display. Half of the trials were 'match' trials, in which all of the studied features were presented again at probe. Half of the trials were 'change' trials, in which one of the studied features was replaced by a new feature on the same dimension. New colors were as equally likely as new shapes. Participants were instructed to ignore any changes in locations or combinations of feature bindings when deciding if a feature was new.

Half of the trials were 'intact binding' trials, in which the color-shape bindings were kept intact between study and probe. The other half of the trials were 'switched binding' trials, in which the color-shape bindings were swapped between all the items in the display. Colors or shapes were swapped with equal probability. On half of the trials ('Cue present'), when the probe display was presented, a cue indicated which item was the target. The target was the item that participants must make a change/no-change decision about. The uncued items never contained a new feature. On the other half of trials ('Cue absent'), no cue was present, and so participants had to make a change/no-change decision about all the items in the display, corresponding to the whole-display condition in Treisman and Zhang (2006). The Cue and Location factors were blocked, and their order counterbalanced. The Feature and Binding factors were randomized within blocks. A schematic of a trial from the two Cue conditions (Absent and Present) is shown in Figure 2. Participants were instructed to respond as accurately as possible. Participants were also instructed to perform articulatory suppression by repeating "Coca-Cola" throughout each trial, to inhibit verbal re-coding. This, and all subsequent experiments, lasted approximately one hour.

Statistical analysis: In each experiment, we applied the following procedure before conducting inferential data analysis. We excluded all trials with RTs above 4,000 ms, or below 100 ms. Corrected hit rates were calculated by subtracting the false alarm rate (the proportion of incorrect feature-change trials) from the hit rate (the proportion of correct feature-match trials). Participants with a mean corrected hit rate less than 0.1 in the baseline condition (Whole-display probe, Old location, Intact binding) were not included in the analysis. However, the qualitative pattern of the data remained the same when all participants were included in the analysis. The reason for removing participants based on their performance in the baseline condition is because performance should be at its highest in this condition. In addition, some of the other conditions produce near-to-chance performance. Therefore, an exclusion criteria based on average performance in all conditions would be too strict, as it would be affected by performance in the close-to-chance conditions. These exclusion criteria were used in all of the reported experiments.

Although we report our conclusions based on null hypothesis significance testing, we have also supplemented our analyses using a Bayesian analysis of variance. The Bayesian approach provides the advantage that it allows one to specify and competitively test null and alternative hypotheses. In contrast to the frequentist approach, where inferences about differences are made on the lack of evidence for a null hypothesis, Bayesian methods provide the relative evidence in favour of the null or alternative hypotheses. Because Bayes factors represent relative evidence between the two hypotheses, indices (BF_{10} for the alternate and BF_{01} for the null) are used to indicate which hypothesis the Bayes factor is describing. For example, a BF_{10} of 9 indicates that there is nine times more evidence for the alternate than for the null hypothesis. Because this value is a ratio, it can also be represented as the amount of evidence for the null: A BF_{10} of 10 is equivalent to a BF_{01} of 0.1. We conducted our analysis using the anovaBF function from the BayesFactor package (Rouder, Morey, Speckman, & Province, 2012) in R (R core team, 2015), using the default JZS prior in all analyses (cf. Bayarri & Garcia-Donato, 2007; Jeffreys, 1961; Rouder et al., 2012; Zellner & Siow, 1980). The method for the following experiments were pre-registered on the Open Science Framework (OSF link: https://osf.io/8ybjp/, https://osf.io/teb6y).

Results

Using our exclusion criteria described above, no participants were excluded from the analysis. However, a total of 92 trials (0.6%) were excluded from the analysis (M = 3, SD = 4.35 trials removed per participant). Mean performance for each condition can be seen in Figure 3. A 2 (Cue: Present vs. Absent) x 2 (Location: Old vs. New) x 2 (Binding: Intact vs. Switched) repeated measures ANOVA was conducted on the mean corrected hit rates. A significant main effect was found for Cue condition [$F\ (1, 31) = 17.39$, $p < .001$, $\eta p^2 = 0.36$, $BF_{10} = 33.35$], in which average performance was 10.24% higher in the Cue absent than in the Cue present. A significant main effect of Location [$F\ (1, 31) = 4.93$, $p = .034$, $\eta p^2 = 0.13$, $BF_{01} = 3.70$], was found, in which changing the locations of the probed items reduced performance by 2.93%. There was a significant main effect of Binding [$F\ (1, 31) = 51.33$, $p < .001$, $\eta p^2 = 0.62$, $BF_{10} = 6.62 \times 10^5$], whereby swapping the feature bindings reduced performance by 13.32%. The two-way Cue x Location interaction was not statistically significant [$F\ (1, 31) = 1.47$, $p = .235$, $BF_{01} = 5.88$]. A significant two-way interaction between Cue and Binding was observed [$F\ (1, 31) = 5.66$, $p = .024$, $\eta p^2 = 0.15$, $BF_{01} = 2.70$], such that changing the bindings in the Cue present condition was more detrimental (Intact: .45, Switched: .28) than in the Cue absent condition (Intact: .49, Switched:.40). Furthermore, the Location x Binding interaction [$F\ (1, 31) = 28.31$, $p < .001$, $\eta p^2 = 0.48$, $BF_{10} = 10.61$] was significant, such that when bindings were kept intact, performance was higher in the old location condition (mean corrected hit rate =.52) than in the New location condition (mean corrected hit rate =.42), whereas when the bindings had switched, performance was lower in the Old location condition (mean corrected hit rate =.32) than in the New location condition (mean corrected hit rate =.36). Finally, the three-way interaction was not significant [$F\ (1, 31) = 0.16$, $p = .694$, $BF_{01} = 7.14$].

Discussion

Experiment 1 extended the study of Treisman and Zhang (2006) by including the use of a cue that indicated which item in the whole probe display was the target. It was expected that if the use of location was obligatory, then we would see the same Location x Binding interaction in both the Cue present and Cue absent conditions. If the use of location was not obligatory and participants could strategically control whether they used location, it was expected that the interaction would only occur in the Cue absent condition. However, the interaction occurred in both Cue present and Cue absent conditions, indicating that the participants employed in-place matching in both conditions. This evidence is indicative that participants did not strategically change their comparison strategy in response to the presence of a cue, which is consistent with the notion that in-place matching may have occurred obligatorily.

Experiment 1 found that when whole-display probes were presented without a cue, changing location or feature bindings reduced performance relative to when locations and bindings remained the same between study and probe displays. This is consistent with previous findings (Hollingworth, 2006, 2007; Jiang et al., 2000; Treisman & Zhang, 2006) and suggests that the change/no-change decision for whole-display probes involves the entire configuration: Comparing each feature at a given location to the feature that was previously at that location, but not comparing it to features at other locations. Furthermore, the same pattern of data was found when the cue was present, in which it could have been possible to conduct a non-spatial search for the cued target. These results alone suggest that exhaustive in-place matching in this task was automatic or obligatory, consistent with existing conceptualizations of VWM (Hayes et al., 2007; Kondo & Saiki, 2012; Papenmeier & Huff, 2014; Silvis & Shapiro, 2014; Treisman & Zhang, 2006).

One potential issue with Experiment 1 was with the presentation timing of the cue. Specifically, the onset of the cue appeared at the same time as the onset of the probe display, and therefore was unlikely to be immediately fixated when the probe display appeared. This may have encouraged a spatial comparison strategy, because in order to decide which item had been cued, the participants potentially needed to attend to each item in turn to make a 'cued/not-cued' decision. If participants needed to examine each object despite the presence of the cue, they may have decided that it was less effortful to make a comparison for each item, rather than visually search for the target and then conduct a memory search once it was identified. Likewise, it is also possible that simply viewing an item will lead to its being automatically encoded in memory, irrespective of whether or not it is task relevant (Olson, Moore & Drowos, 2008), and therefore making a 'task relevancy' decision about each item may have led to them all being encoded.

A further possible issue with the type of cue used in this experiment, which was a square outline surrounding the target item presented during the test display, is that it may produce interference with the presentation of the probed items, via visual masking (Breitmeyer & Ganz, 1976), which might explain why performance was slightly lower in the Cue present condition.. It is also possible that the presentation of the cue biased participants towards responding change, because the presentation of the cue is a distinctive change between the study and probe displays. In light of these considerations, Experiment 2 replicated Experiment 1, with the exception that the probe display cue was removed and replaced with a cue presented at the location of the target during the maintenance interval.

Experiment 2

Experiment 2 replicated Experiment 1, with the exception that the cue was replaced by a cue presented during the maintenance interval at the location of the target item. The duration of the cue's presentation time was chosen to direct covert, but not overt, attention (Carrassco & McElree, 2001; Jonides, 1981).

Method

Participants. Thirty-one participants (aged 17-46, 25 were female) participated in the study.

Materials, design and procedure: The materials, stimuli, design, and procedure of Experiment 2 were identical to that of Experiment 1, with the following exceptions. A cue consisting of a white dot (visual angle: 0.9° x 0.9°) was presented during the maintenance period at the location of the target item. The cue was displayed for 67 ms, and its offset occurred 50 ms prior to the onset of the probe display. These timings were selected because they have been shown to shift endogenous, but not exogenous, attention (Carrassco & McElree, 2001; Jonides, 1981). Shifting endogenous attention was important because it ensured that the participants knew the location of the target item before the on-set of the test display. At the same time, the cue-probe asynchrony was short enough that participants could not saccade to its location before the test display, ensuring that all items were displayed before the participant moved overt attention toward the target. No cue was present during the probe display. A schematic of the procedure is presented in Figure 4.

Results

Using the same exclusion criteria as Experiment 1, five participants were excluded from the analysis (although the pattern of results remained the same if all participants were included in the analysis). Of the remaining participants, a total of 106 trials (0.65%) were excluded (M = 4.07, SD = 7.74 trials removed per participant). The mean corrected hit rates for each condition are presented in Figure 5. A 2 (Cue: Present vs. Absent) x 2 (Location: Old vs. New) x 2 (Binding: Intact vs. Switched) repeated measures ANOVA was conducted on the mean corrected hit rates. The main effects of Cue [F (1, 25) = 1.85, *p* = .186, BF_{01} = 3.34] and Location [F (1, 25), *p* = .264, BF_{01} = 4.54] were not significant. However, the main effect of Binding was significant [F (1, 25) = 25.48, *p* < .001, ηp^2 = 0.51, BF_{10} = 30.31], whereby swapping the bindings reduced average performance by 8.42% (Intact: .38, Switched: .31 corrected hit rates). The Cue x Location interaction [F (1, 25) = 0.65, *p* = .427, BF_{01} = 6.25], and Cue x Binding interactions [F (1, 25) = 1.77, *p* = .195, BF_{01} = 5.55] were not significant. The two-way Location x Binding interaction was significant [F (1, 25) = 27.65, p < .001, ηp^2 = 0.54, BF_{10} = 3.19]. Specifically, when Location was old, performance was higher in the Old location (mean corrected hit rate = .43) than in the New location condition (mean corrected hit rate = .34). However, when the bindings had swapped, performance was lower in the Old location condition (mean corrected hit rate = .28) than in the New location condition (mean corrected hit rate = .32). Critically, the three-way interaction between Cue, Location, and Binding was not statistically significant [F (1, 25) = 1.87, p = .184, BF_{01} = 4.90].

Discussion

In Experiment 2, a Location by Binding interaction was observed, irrespective of whether the cue was presented or not. These results, like Experiment 1, are consistent with the idea that participants are automatically using location to match the probe display with the contents of their memory.

Experiment 3

The results of Experiment 2 indicate that there did not appear to be an effect of Cue, beyond the finding that presenting the cue made performance numerically worse (in Experiment 1, this decrease was statistically significant). It is important to rule out the possibility that participants may have missed the cue on a substantial number of trials, given it is presented peripherally and for a very brief duration. If the participants were not able to detect the cue on a trial, then that trial would be perceived as a Cue absent trial, and participants would be forced to use in-place matching, weakening the opportunity to observe any effect of cue. Therefore, Experiment 3 included a measure of participants' ability to detect the cue. This was achieved by probing participants as to the location of the cue on a subset of trials, and therefore the extent to which they were potentially making a decision on the basis of a single item versus the entire display.

In addition, we introduced a feature change amongst the uncued items on 50% of trials. This latter manipulation allowed us to measure the extent to which the uncued items influenced the match/change decision. If participants are performing exhaustive in-place matching, the features of uncued objects will affect the detection decision that is made: in particular, changes at uncued locations in the Old location condition should make people more willing to make a "change" response. In contrast, if decision-making is based solely on the cued item, then the nature of uncued items should not affect performance, so that accuracy for trials including a changed uncued item should be identical to trials in which the uncued items did not change.

Method

Participants. Thirty-two naïve participants (aged 18-24, 25 were female) participated in the study.

Materials, design and procedure: The materials, stimuli, design, and procedure of Experiment 3 was identical to that of Experiment 2 with the following exceptions. First, only intact bindings were presented, and switched binding trials were excluded, as the Location x Binding interaction was not central to the question of this experiment. Second, 12.5% (give frequency per condition) of the trials were followed by catch trials. In these trials, the probe display was re-presented with an instruction in the top left of the screen to click on the location of the cued item. Participants made their response using the mouse, and could select any location on the screen, including locations where stimuli were not presented. Third, on 50% of trials, one of the uncued items would contain a new feature, with equal probability of it being a shape or color. Participants were instructed to ignore these changes, and only decide about the cued item. Finally, because irrelevant feature changes could occur, all trials were Cue present trials: If the cue was not presented, it would not be possible to distinguish between a ‘distractor’ feature change, and a ‘target’ feature change. Therefore, it was necessary to always present the cue so that participants could make this distinction.

Examples of probe displays with combinations of target match/change, crossed with distractor match/change, are presented in Figure 6. In the left two probe display examples of Figure 6, participants should respond ‘match’, because the target, the red square, does not contain a new feature. This is even the case in the bottom left example, where an uncued item, the blue circle, has changed color to purple. Likewise, in the right two panels, participants should respond change, because the target contained a new feature: the square is now green, irrespective of the status of the uncued items.

Results

Using the same exclusion criteria described in Experiment 1, a total of 153 trials (0.93 %) were excluded (M = 4.78, SD = 8.70 per participant). We analyzed proportion of change responses separately for match and change trials so as to examine the independent effects of target and distractor changes on performance. Figure 7 plots the proportion of change responses as a function of target, distractor and location changes. We conducted a 2 (Target: Match vs. Change) x 2 (Distractor: Match vs. Change) x 2 (Location: Match vs. Change) within-subjects ANOVA on mean proportion of change responses. The main effects of Target [F (1, 31) = 219.52, $p <$ .00, ηp^2 = .87, BF_{10} = 3.83 x 10^{41}], and Distractor [F (1, 31) = 70.35, $p < .001$, ηp^2 = .69, BF_{10} = 85.98 x 10^{3}] were significant. Participants responded change more often when a target had changed (.72) than when no target had changed (.38). Likewise, they were more likely to respond change when a distractor had change (.62) than when it had not (.47). The main effect of Location was not significant [F (1, 31) = 3.18, p = .084, BF_{01} = 5.31]. The two-way interaction between Target and Distractor was significant [F (1, 31) = 31.16, $p < .001$, ηp^2 = 0.51, BF_{10} = 2.09]. Specifically, when the targets matched, a change in the distractors increased the proportion of change response by 22.19 percentage points. However, when the target had changed, a change in the distractors increased the proportion of change responses by 8.12 percentage points. The two-way interaction between Target and Location was statistically significance [F (1, 31) = 27.34, $p < .001$, ηp^2 = .47, BF_{01} = 3.788]. The two-way Distractor x Location interaction was not significant [F (1, 31) = 0.04, p = .85, BF_{01} = 7.29]. Finally, the three-way interaction between Target, Distractor, and Location, was significant [F (1, 31) = 11.57, p = .002, ηp^2 = .31, BF_{01} = 5.41].

In order to examine participants' ability to determine which item was the target, the experiment measured memory for the cued target's location (this test being presented on 12.5% of the trials). For this analysis, a correct response was coded as correct when the *x* and *y* co-ordinates of the mouse-click fell within a square subtending 1.8° x 1.8° of visual angle around the centre of the target. Figure 8 shows mean accuracy by participant. Most participants average cue detection rate was above chance (Figure 8). However, there was considerable variability in detection rates, with some participants appearing to perform near, or below, chance. The dashed line of Figure 8 represents chance performance – the probability of selecting the target location, if they randomly selected one of the three possible probe item locations. However, it was possible for participants to select any location on the screen.

One question of interest is how cue detection performance is related to performance on the recognition task. To answer this question, we conducted exploratory analyses by calculating the correlation between performance in the catch-trial task and that in the change detection task. The means for individual participants, and lines fit to those data are presented in Figure 9. Target sensitivity was measured as the mean difference between target match and target change trials. Distractor sensitivity was measured as the mean difference between distractor match and distractor change trials. When the distractors matched, there was no correlation between catch trial performance and target sensitivity [r (30) = -.14, p = .447, BF_{01} = 2.42] (Figure 9a). However, there was a significant positive correlation with target sensitivity when the distractors changed [r (30) = .47, p = .006, BF_{10} = 7.11] (Figure 9B). We also found a significant negative relationship between catch trial performance and distractor sensitivity when the targets matched [r (30) = -.72, p < .001, BF_{10} = 3739.19] (Figure 9c) and when the targets had changed [r (30) = .36, p = .043, BF_{10} = 1.74] (Figure 9d). That is, participants who were less able to detect the cue were more likely to detect changes on the basis of distractor features. There was no significant correlation [r (30) = .31, p = .08, BF_{10} = 1.06] between catch trial performance and location change sensitivity, measured by the mean difference between the Old location trials and New location trials (Figure 9e).

Discussion

Probing participants about their memory of the cued target location on a subset of trials allowed us to determine whether participants did not detect the cue, or were choosing not to utilize it in their change detection decision. There was substantial variability between participants in their ability to detect the cue, with some participants performing near, or below, chance. These data suggest that some participants were unable to detect the cue (or if they did detect it, they appeared not to maintain a memory of its location across the duration of the trial). Therefore, the most likely explanation for the lack of apparent difference between Cue present and Cue absent conditions in Experiment 2 is that a third of participants were simply not able to detect the cue, whilst the remaining two-thirds could detect the cue, but did not make use of it in order to employ a non-spatial comparison process.

The second modification in Experiment 3 was the introduction of task-irrelevant changes to the uncued (distractor) items, which participants were instructed to ignore. If participants were using only the cue to inform their responses, then performance should be unaffected by these irrelevant changes, because the cue can be used to filter-out the irrelevant items from the decision or comparison stage. The significant main effect of distractor feature (match or change) showed that when a distractor contained a new feature, participants were more likely to respond 'change'. Therefore, at least on some trials, participants did not utilize the cue, and erroneously incorporated information about the distractors into their final match/change decision. This interpretation is further supported by the finding that there was a significant effect of changing item locations: If participants used the cue, they could have employed a non-spatial search of memory for a match, in which case, they would have been unaffected by changes in item locations. So far, the data is consistent with the idea that participants are unable to flexibly switch between different comparison strategies and that a configuration-based comparison process is automatic.

Finally, we conducted exploratory analyses by correlating individual performance on the catch trials against individual performance in the change detection task. These interpretations are made with caution because the analysis was exploratory, not confirmatory, and limited by the relatively small sample size for an individual differences analysis. Because the data set contained a mix of participants who could detect the cue to varying extents, it is possible that those who could detect the cue were using a different comparison strategy to those who could not detect the cue. The data showed that participants with lower performance on the catch trials were more severely affected by distractor changes than those participants who performed well in the catch trials. Furthermore, when the distractors had not changed, both groups were equally effective at detecting a change in the target item. These data are consistent with the idea that all participants employed a spatial comparison process, but that they varied in the extent to which information about the distractors entered the decision stage. For example, the participants who did not detect the cue would not be able to distinguish between a target change and a distractor change. As a result, they may have used global in-place matching, and would therefore respond ‘change’, irrespective of whether they detected a target change or a distractor change. In contrast, the participants who could detect the cue may have used the cue, and the spatial configuration, in order to identify the target item, and make an in-place comparison for that item only. This opens the possibility of a strategic switch between global in-place matching and partial in-place matching, whereby the configuration is utilized for a spatial comparison of the target item.

Experiment 4

One possible explanation for the results of Experiment 2 was that participants were not able to utilize the cue, possibly due to its brief peripheral presentation. The results of Experiment 3 gave some credence to this possibility, finding that only some participants could reliably identify the location of the cued item. It is also possible that participants were able to detect the cue, but the change detection test display retroactively interfered with their memory for the cued-location. Finally, it is also possible that they were unable to detect the cue at all, either because the cue’s presentation was too short for them to detect, or because detecting the cue was so effortful that they chose not to encode or report its location accurately. Experiment 4 addressed these issues by presenting a more salient cue that appeared in a predictable central location. The cue in Experiment 4 appeared during the maintenance interval, for a longer period of time than in Experiments 2 and 3. The cue was a large, centrally presented arrow pointing to the location where the target would subsequently appear in the probe display. These changes were made to ensure that participants could make strategic use of the cue in the change detection task.

The design of Experiment 4 was the same design as in Experiments 1 and 2. Changes in the locations of items and bindings between features could occur, and participants were instructed to ignore both. Participants were told to decide if a new feature was presented, and that if there was a new feature, it was always a 'valid' new feature, and that 'invalid' new features in the uncued items never occurred. Because the uncued items never contained new features, the uncued trials were re-introduced, making it possible to compare patterns of data across Cue present and Cue absent trials. Additionally, although this is a replication of Experiment 2, we also kept the catch trials from Experiment 3, in which participants were probed about the location of the cued item. As with Experiment 3, the addition of the catch trials allowed us to measure the extent to which participants were detecting the cue, and therefore make inferences about their decision processes in the change detection task. If they missed, or did not make use of the cue, they would not know which item was the target, and therefore would have no choice but to employ exhaustive in-place matching. However, if they could reliably use the cue, participants would have the option of using a comparison strategy other than in-place matching. Thus, by measuring the cue detection rate, it would be possible to rule out the explanation that the use of in-place matching was simply due to an inability to use the cue. As a measure of the rate with which participants used the cue, we again probed the cue location, this time after 25% of the cued trials. We increased the proportion of catch trials because the relative number of cued trials was halved; in order to maintain the same total number of catch trials, the proportion needed to be doubled.

Finally, we conducted replications of Experiment 4 and 5 with larger sample sizes because we had an opportunity to test many participants at once. The methods of each replication match those of the original study. However, because we conducted the replications on a larger sample, with data being collected from multiple participants simultaneously (a PC was assigned to each participant) as part of a laboratory class, the replications of Experiment 4 and 5 did not employ articulatory suppression. The results of the replication will be presented alongside the analysis of the original study. Our interpretation of the data was based on the analysis of the original run of the experiment. However, our interpretations are bolstered by the replications.

Method

Participants. Thirty-two naïve participants (aged 18-36, 24 women) participated in the one-hour study participated for course credits. The replication was conducted on seventy-eight participants (aged 18-30, gender data were not collected).

Materials, design and procedure: The materials, stimuli, design, and procedure of Experiment 4 were identical to that of Experiment 2, with the following exceptions. The cue took the form of an arrow, 1.8^{o} x 1.8^{o} of visual angle, presented in the center of the screen during the maintenance period. The orientation of the arrow changed, so that the tip pointed towards the location where the target item would subsequently appear in the probe display. The cue duration was changed to 200 ms during the interval, and off-set 100 ms prior to the on-set of the probe display. The central location was never used to present an item. A schematic of the procedure is presented in Figure 10. After 25% of the Cue present trials the cues location was also probed, in which the probe display was re-presented and participants were asked to indicate the location of the target item.

The participants in Experiment 4 used articulatory suppression throughout the task, however the participants in the replication did not, as they were tested in groups. Additionally, due to a programming error, the replicated experiment did not contain cue-location probe trials.

Results

Five participants were excluded from the analysis following application of the exclusion criteria outlined in Experiment 1. Of the remaining participants, a total of 89 trials (0.54%) were excluded (M = 3.29, SD = 4.46 trials removed per participant). The mean corrected hit rates for each condition are presented in Figure 11. A 2 (Cue: Present vs. Absent) x 2 (Location: Old vs. New) x 2 (Binding: Intact vs. Switched) repeated measures ANOVA was conducted on the mean corrected hit rates. The main effect of Cue was significant [F (1, 26) = 11.98, p = .002, ηp^2 = 0.31, BF_{10} = 2.93]. The main effect of Location was not significant [F (1, 26) = 2.68, p = .114, BF_{01} = 4.00]. The main effect of Binding was significant [F (1, 26) = 69.28, p < .001, ηp^2 = 0.73, BF_{10} = 40824.21], whereby changing the Bindings reduced performance by 13.8%, relative to keeping the Bindings intact. The two-way interaction between Cue and Location was not significant [F (1, 26) = 3.66, p = .067, BF_{01} = 3.57]. However, the two-way interaction between Cue and Binding was significant [F (1, 26) = 20.27, p <.001, ηp^2 = 0.43, BF_{10} = 1.32], whereby the effect of Binding in the cue condition was less (19.7%) than that of the Binding effect in the Cue absent condition (7.8%). The Location x Binding interaction was significant [F (1, 26) = 17.21, p <.001, ηp^2 = 0.4, BF_{10} = 5.27], whereby changing locations when Binding were intact had a detrimental effect (Old location: 0.45, New location: 0.34), whereas changing locations when Bindings were switched had a beneficial effect on performance (Old location: 0.24, New location: 0.28). Finally, the three-way Cue x Location x Binding interaction failed to reach statistical significance [F (1, 26) = 1.90, p = .18, BF_{01} = 5.26].

The same analysis was conducted on the data collected in the replication experiment. Using our exclusion criteria, six participants were removed, and a total of 1,202 trials (3.26%) were excluded (M = 16.69, SD = 30.41 trials removed per participant). The means for the replication of Experiment 4 are presented at the bottom of Figure 11. There was a significant main effect of Cue [F (1, 71) = 33.32, p < .001, ηp^2 = 0.32, BF_{10}], Location [F (1, 71) = 15.97, p < .001, ηp^2 = 0.18, BF_{10} = 1.93], and Binding [F (1,71) = 114.52, p < .001, ηp^2 = 0.61, BF_{10} = 2.3×10^{12}]. The Cue x Location interaction was not significant [F (1, 71) = 3.5, p = .066, BF_{01} = 5.36]. The Location x Binding interaction [F (1, 71) = 40.09, p < .001, ηp^2 = 0.36, BF_{10} = 53.36], Cue x Binding [F (1, 71) = 14.98, p < .001, ηp^2 = 0.17, BF_{10} = 1.62], and the Location x Binding x Cue interaction [F (1, 71) = 7.39, p = .008, ηp^2 = 0.09, BF_{01} = 3.99] were statistically significant. Finally, we assessed whether there were differences in outcomes for the experiment and replication by conducting a Cue x Location x Binding x Experiment mixed ANOVA. There were no statistical differences in performance between the two experiments (all p's > .05, Fs < 1 for the Experiment factor).

Figure 12 shows performance for each participant in the catch trial task in Experiment 4. The dashed line represents chance performance – the probability of selecting the target location, if they randomly selected one of the three possible probe item locations. It was possible for participants to select any location on the screen, and empty screen locations were selected 11% of the time. Although the majority of participants performed well, we identified some participants as being unable to accurately report the location of the cue. Experiment 4 used a more salient cue than Experiment 3 and the average detection rate of the cue was higher in

Experiment 4 (M = 0.77, SD = 0.28) than in Experiment 3 (M = 0.71, SD = 0.27), although this difference failed to reach statistical significance [t(57) = 0.38, *p* = .389].

Discussion

The results of Experiment 4 replicate the findings of Experiments 1 and 2. Specifically, there was a Location x Binding interaction, suggesting that participants were using item locations to guide their comparisons. Furthermore, this interaction was found in both the Cue absent and Cue present conditions, despite the high accuracy in responding to the cue location task. In the Cue present condition, participants who detected the cue had the option to switch to a non-spatial comparison strategy, analogous to when memory is only probed for a single item. However, it appears that they did not utilize this option and continued to use in-place matching. This finding is consistent with the idea that in-place matching is automatic and that participants are unable to switch between comparison strategies. However, taking into account the results of Experiment 3, in which those who could detect the cue were less sensitive to distractor feature changes, and the fact that the cue was highly detectable in Experiment 4, a second possibility is that the high performing detectors were using the spatial configuration to guide recognition, but did not take the features of the distractors into account when making their decision. In other words, those participants may have used partial, rather than global, in-place matching, in which they made use of the spatial configuration of the display in order to identify the location of the target and use a spatial comparison of the target. This possibility is explored further in Experiment 5.

In addition to the pre-registered experiment, we conducted a direct replication of Experiment 4, with the exceptions that multiple participants were tested in the same lab at once, and were not doing articulatory suppression during the task. The results of the replication are highly consistent with the results of the original study. Because we found the same result in experiments with and without articulatory suppression, these results indicate that effects of VWM in this particular design can be detected without the need for articulatory suppression. However, because of the very short timings of the stimuli in this particular task, it may not have been possible to sub-vocally rehearse in this task at all, even when participants were not using articulatory suppression. It is possible that articulatory suppression may have a more beneficial effect when the stimuli are presented for longer periods of time. Interestingly the Location x Binding x Cue interaction was statistically significant in the replication, but not in the original version of the experiment. However, there was very little difference in the qualitative pattern of the data between the two experiments, and the Bayes factor suggested that there was only about four times as much evidence in favour of there being an interaction, than there being no interaction (BF10 = 3.99). It is possible that the significant interaction was driven by higher performance in the 'Intact, Old location' condition in the Cue present condition, but not in the Cue absent condition. From a visual inspection of the data, it does not appear that participants were employing a non-spatial comparison strategy in the Cue present condition in either experiment.

One important feature of Experiment 4 was that the cue was easily detectable. Some participants still performed poorly in the catch trials, despite the highly salient cue. Because of the high saliency of the cue, it seems more likely that the participants who performed poorly at identifying the location of the cue did so because they were strategically ignoring the cue, in order to minimize effort, rather than them being unable to detect the cue at all. However, we should add a caveat that a subset of those participants may be making use of the cue, but either unable to recall the cues location, or chose not to respond accurately as to its location in the catch trials. A further possible caveat is that the change detection probe, which occurs between the presentation of the cue and the probe of the cues location, may have served as strong retroactive interference for the cue's location. This explanation seems unlikely, considering the large proportion of participants who were able to correctly recall the cue's location.

Experiment 5

In Experiments 1-4 at least some, if not all, participants appeared to use in-place matching. Those participants who did not detect the cue seemed to have used global in-place matching, comparing each item in the test display to the stored item at the same location. The participants who had high detection rates for the cue appeared not to use global in-place matching (based on the results of Experiment 3), but appear to have at least used a partial in-place matching, a spatial comparison that incorporated the configuration of the distractor items, as those participants showed the same Location x Binding interaction for Cue absent and Cue present trials. Participants appeared to use these spatial comparison processes despite the fact that provision of a cue should have reduced the need to use a spatial comparison process. These findings suggest that although global in-place may not be automatic, our participants consistently used a spatial matching strategy (either global- or partial- in-place matching) in order to perform change detection.

One outstanding issue with Experiments 1, 2, and 4, is that although the design gave the opportunity for participants to use either a spatial or non-spatial comparison strategy, the task may not have sufficiently encouraged a non-spatial strategy. Specifically, a feature-change was always a 'valid' feature change, in that only the target item ever contained a new feature, and non-target items always contained old features. As a result, conducting in-place matching should be able to detect the change as effectively as a non-spatial comparison process would, despite different decision rules. For example, if exhaustive in-place matching is used, each item is checked for a change and the participants respond change on the basis of whether at least one of the comparisons is a miss-match irrespective of whether or not the change was the cued item (Figure 1a). In contrast, participants could employ the strategy of non-spatially searching memory for the target (Figure 1c and 1d). In this case, they simply need to respond match if a match is detected during the search, and respond change if they have exhaustively searched memory but failed to find a match. Both of these comparison strategies should be equally effective at detecting a change in the target, however a non-spatial comparison is arguably more effortful, because it might depend on actively making use of the cue and inhibiting the comparisons of uncued items. If there was no strategic benefit for using one strategy over another, then participants might simply use the least effortful one, in this case being exhaustive in-place matching.

This strategic explanation for the lack of three way interaction between Cue, Location, and Binding in Experiments 1, 2, and 4 may also provide an explanation as to why the participants who could recall the cued location in Experiment 3 were less sensitive to distractor changes than those who could not recall the cued location. In Experiment 3, irrelevant distractor changes did occur, which participants were instructed to ignore. The distractor changes may have acted as a disincentive to use global in-place matching, a comparison strategy which does not distinguish between relevant and irrelevant changes. Instead, the participants who could detect the cue in Experiment 3 appear to have used a different kind of comparison strategy, one that uses the global configuration to guide the spatial comparison of the target item, but one in which the distractor items do not enter the decision stage. Therefore, one important factor affecting the strategic choice of one comparison strategy over another seems to be how a particular comparison strategy can deal with irrelevant changes, and whether or not irrelevant changes are likely to occur on a trial.

Therefore, Experiment 5 aimed to make global in-place matching costly---in terms of performance---by increasing the probability that an uncued item would change. If there is a high probability that an uncued item has changed, global in-place matching would be a poorer strategy choice compared to situations where no (Experiments 1 and 2) or few (Experiment 3) irrelevant changes were likely to occur. If participants ignore the cue, and conduct exhaustive in-place matching—whereby the decision rule is to respond change in response to any change—there would be a higher probability of erroneously responding ‘change’ to an invalid change in one of the uncued items. Thus, relying on the cue, and conducting a non-spatial search, should be encouraged because it avoids increasing false alarms caused by the uncued changes.

Experiment 5 replicated Experiment 4 with the exception that there was a 75% probability on each trial that one of the uncued items contained a new feature. Participants were instructed to ignore such changes. Because of this manipulation, all trials had the target cued, because the target versus distractor distinction does not apply to uncued trials.

Finally, we also conducted a direct replication of Experiment 5, using a larger sample size. The method of the replication matches that of Experiment 5, with the exception that multiple participants were tested in the lab at once, and as a result, were not employing articulatory suppression during the task. The results of both the original experiment and replication study are described below.

Method

Participants. Thirty-two participants (aged 18-32, 19 were female) participated in the study. In the direct replication, we tested 79 new participants (aged 17-22, 60 female).

Materials, design, and procedure: The materials, stimuli, design, and procedure of Experiment 5 was identical to that of Experiment 4, with the following exceptions. On 75% of trials, one of the uncued items contained a new feature, with equal probability of it being a new color or new shape. In addition, Cue absent trials were excluded, and the number of Cue present trials was doubled over the number in Experiment 4. The same central cue was used as in Experiment 4. We also conducted a replication of Experiment 5 with a larger sample size and without articulatory suppression during the task.

Results

Seven participants were excluded from the analysis based on the exclusion criteria. From the remaining participants a total of 122 trials (0.74%) were excluded (M = 4.88, SD = 6.15 trials removed per participant). Mean corrected hit rates in each of the Location and Binding conditions are presented in Figure 13. Additionally, Figure 14 shows mean proportion change across the different Target, Distractor, and Location conditions. First, we excluded trials in which a distractor had changed, to examine the main focus of this experiment: the interaction between Location and Binding. A 2 (Location: Old vs. New) x 2 (Binding: Intact vs. Switched) repeated measures ANOVA was conducted on corrected hit rates. The main effects of Location [$F(1, 24) = 0.70$, $p = .411$, $BF_{01} = 5.55$] and Binding [$F(1, 24) = 0.75$, $p = .394$, $BF_{01} = 6.25$] were not significant. The Location x Binding interaction was not significant [$F(1, 24) = 0.23$, $p = .639$, $BF_{01} = 4.76$]. Next, we examined the effects of target and distractor change on the proportion of change responses. Switched Binding trials were removed from this analysis. A 2 (Target: Match vs. Change) x 2 (Distractor: Match vs. Change) x 2 (Location: Match vs. Change) repeated measures ANOVA was conducted on the proportion of change responses in each condition.

The main effect of target feature [F (1, 25) = 88.15, $p < .001$, $\eta p^2 = .79$, $BF_{10} = 3.89 \times 10^{43}$], distractor feature [F (1, 24) = 31.19, $p < .001$, $\eta p^2 = .57$, $BF_{10} = 68.58$], and Location [F (1, 24) = 6.56, $p = .017$, $\eta p^2 = .2$, $BF_{01} = 3.77$] were significant. The mean proportion change for target change trials was 59.25%, relative to 33.9% on target match trials. A change in the distractor resulted in an increase in proportion change responses, from 42.78% on distractor match trials, to 50.38% on distractor change trials. The mean proportion of change responses was higher on New location trials (47.47%) than on Old location trials (45.69%).

The two-way interaction between target and distractor was also significant [F (1, 24) = 28.38, $p < .001$, $\eta p^2 = .55$ $BF_{10} = 1.93$]. When the targets matched, a change in the distractor increased the proportion of change responses by 11.17 percentage points, from 28.84% to 40.56%. However, when the target had changed, a change in the distractors only increased the proportion of change responses by 2.31 percentage points, from 60.37%, to 58.06%. The two-way interaction between Distractor and Location [F (1, 24) = 4.19, $p = .052$, $BF_{01} = 4.25$] and the three-way Target, Distractor, and Location interaction [F (1, 24) = 1.35, $p = .257$, $BF_{01} = 7.65$] were not significant.

The same analysis was also conducted on the data from the replication study. Thirteen participants were excluded from the analysis due to low performance. Of the remaining participants, 1,025 trials (M = 15.53, SD = 33.88) were removed. There was a significant main effect of Location [F (1, 65) = 35.89, $p < .001$, $\eta p^2 = 0.36$, $BF_{10} = 3.19\text{x}10^4$), Binding [F (1, 65) = 70.36, $p < .001$, $\eta p^2 = 0.52$, $BF_{10} = 7.21\text{x}10^5$), and a significant Location x Binding interaction

[F(1, 65) = 23.88, $p < .001$, $\eta p^2 = 0.27$, $BF_{10} = 2.09$). In contrast to the original version of the experiment, there were significant effects of Location and Binding, as a well as a significant two-way interaction. However, the direction of these effects occurred in the opposite direction to that predicted by in-place matching, such that participants were more accurate when locations had changed, than when they had not changed. The means for each condition in the replication can be found in Figure 13 and Figure 14.

Discussion

Experiment 5 was designed to strongly discourage participants from using global in-place matching by greatly increasing the probability of an uncued item changing, and hence inflating the false alarm rate for participants using this comparison strategy. The results of Experiment 5 showed that there was no interaction between the Location and Binding factors. This finding suggests that, in this experiment, participants utilized the cue and employed a non-spatial comparison. These conclusions are further supported by the replication of Experiment 5 that largely replicated the pattern of results. In fact, in the replication, we found that changing item locations improved accuracy, rather than reduced it, as would be predicted if the participants were employing in-place matching. One issue with this interpretation is that there was still a detrimental effect of changing the features of the uncued items. If participants were conducting a non-spatial search for only the cued target item (akin to non-spatial search of single probes; e.g., Gilchrist & Cowan, 2014; Treisman & Zhang, 2006) then they should have been unaffected by these task-irrelevant feature changes. We address this issue and provide some plausible explanations for this effect in the General Discussion.

General discussion

The aim of this study was to assess the flexibility of VWM, in terms of the ability to use different comparison strategies under different task demands. Many view the structure of VWM as fundamentally spatially organized (e.g., Golomb et al. 2014) such that non-spatial visual information, such as color and shape, is necessarily bound to its location (Golomb et al. 2014; Eimer & Grubert, 2014; Hollingworth & Rasmussen, 2010; Hollingworth, 2007; Jiang et al. 2000; Kondo & Saiki, 2012; Olviers & Schreij, 2014; Pertzov & Husain, 2014; Rajsic & Wilson, 2014; Silvis & Shapiro, 2014; Treisman & Zhang 2006; Vidal et al. 2005). Despite this viewpoint, VWM has also classically been viewed as a system that allows for complex cognition, such as reasoning and novel problem solving (Baddeley & Hitch, 1974; Hitch, 1978). To achieve this flexibility, VWM would require an executive process capable of flexible reconfiguration according to current goals (Oberauer, 2009). Automatically binding features to locations may limit our ability to flexibly respond to situations where location is not a diagnostic feature of the task. For example, to search for a friend amongst a crowd, a template of the friend stored in VWM (Bundensen, 1990; Desimon & Duncan, 1995) may be compared to multiple different individuals, each occupying different locations in the environment. However, spatial information may be highly diagnostic in other tasks - for example, when re-stablishing object correspondence after disruptions caused by occlusions and saccades (Hollingworth & Franconerri, 2009). Therefore, a truly flexible VWM system would require the ability to select the most appropriate comparison process for the particular task at hand, which may not necessarily always rely on location.

In the introduction, we outlined the five assumptions of exhaustive in place matching. These assumptions were i) task-relevant visual features are automatically bound to their spatial locations during encoding; ii) each probed item is automatically compared to the memory representation at the corresponding location; iii) probed items are not compared to memory representations at non-corresponding locations ; iv) this comparison process occurs for every item presented in the test display; and v) the evidence from all of these comparisons is integrated in order to form a decision, such as whether or not a change has occurred in the display. We will now evaluate each of these assumptions below.

The first assumption (obligatory location encoding) could not be directly addressed by this set of experiments because we were interested in the processes during the comparison and decision stages. However, we can conclude that locations were encoded in Experiments 1-4, because of the detrimental effect of changing locations on performance in those experiments. However, because location did not appear to enter in the decision process in Experiment 5, there was no way to measure whether locations were encoded in the study displays of Experiment 5. It is possible that locations were obligatorily encoded, but not relied upon during the decision stage, or that because participants did not expect to use location in the decision, they did not encode it in the first place (e.g., Vincente-Grabovetsky et al. 2012). Regarding the first assumption that locations are automatically encoded, our results leave open the possibility that it may be possible to encode many different aspects of a scene, including locations, but that these relations are not always relied upon during the comparison and decision stages. For Experiments 1-4, the finding that changing locations makes performance worse indicates that the probed items were compared to the memory representations bound to the corresponding locations (assumption ii). However, we have also found conditions (Experiment 5) in which changing item locations did not have a detrimental effect on performance and have therefore shown that this assumption may not always hold true. Additionally, in Experiments 1, 2, and 4, we found that when locations were kept the same, but the feature bindings had switched, performance was substantially worse than when the bindings had not switched. These findings support the assumption that the participants did not compare probed items to memory representations at other locations (assumption iii), because it would have appeared to the participants as if each location contained a ‘new’ feature, which was in fact presented elsewhere in the display. However, Experiment 5 did not replicate this finding, suggesting that the participants were in fact able to compare the cued item to multiple memory representations at separate locations (refuting assumption iii). Furthermore, obligatory in-place matching is thought to occur for all of the probed items (assumption iv). This appears to be the case in Experiment 3 and 5, in which changes in the to-be-ignored items increased participants’ bias towards responding ‘change’, indicating that people were using information from the uncued items in making their match decision, despite having been explicitly instructed not to do so. However, Experiment 3 also showed that when participants were able to report the location of the cued item, they were also able to ignore irrelevant distractor changes, refuting assumption iv. Finally, our experiments do not directly address assumption v (evidence is integrated to form a response decision). Future research is needed to fully understand how decision rules change in response to strategic changes in the comparison process.

Taken together, these results suggest that global in-place matching is one comparison strategy which is often employed in the change detection task and that there are conditions under which the assumptions of global in-place matching hold true (e.g., Experiments 1, 2, and 4). However, the findings indicate that there are also conditions where these assumptions do not hold true (e.g. Experiments 3 & 5) suggesting that global in-place matching is not strictly an obligatory process and that other types of comparison strategies can be employed.

Our data suggest that participants can flexibly alternate between different comparison strategies in the whole-display probe change detection task. Specifically, we suggest that participants can either employ a spatial comparison process, which either compares all items in the display (exhaustive in-place matching: Experiments 1, 2, and 4) or only a sub-set of relevant items (partial in-place matching: Experiment 3). When the task demands strongly discourage any form of in-place matching, a non-spatial comparison can be conducted, in which memory is searched for a match with a single target item (Experiment 5), akin to the pattern observed when only a single item is presented as a probe (Donkin & Nosofsky, 2012; Gilchrist & Cowan, 2014; Sternberg, 1966). The finding from Experiment 5 supports the notion that individuals can flexibly employ different comparison strategies in response to different task demands, allowing for more flexible higher-level cognition and problem solving.

One finding of particular interest was that, for Experiment 2, the numerical difference between old and new locations was larger in the Cue absent condition than in the Cue present condition (Figure 5). Although this interaction was not statistically significant, this is the pattern of data expected if participants were to conduct in-place matching on Cue absent trials, but a non-spatial search on Cue present trials. It is possible that the lack of a statistical interaction could be explained by a sub-set of participants who failed to detect the cue, and were therefore forced to employ in-place matching on both Cue absent and present trials. Unfortunately, because we did not measure the rate of cue detection in Experiment 2, it is not possible to determine how many participants were able to detect the cue. However, in Experiment 3, we used the same type of cue as in Experiment 2 on a sample from the same population, in which we did measure the rate of cue detection. The vast majority (90.65%) of participants performed above chance in the cue-location probe in Experiment 3, which gives us some confidence that the participants could accurately detect the cue in Experiment 2. An additional finding of note was that this pattern of data was reversed in Experiment 4. Specifically, there was a much larger numerical difference between Old location and New location trials on the Cue Present trials, than on Cue Absent trials (Figure 11). This comes a surprise, as it is the reverse of what we expected if participants were changing their response strategies. Again, this interaction was not statistically significant, and therefore it is not possible to draw any strong conclusions. However, the poorer performance on Cue present, New location trials, may have resulted from confusion about how to use the cue. Specifically, on old location trials, the cue pointed to the location of a previously presented item, and therefore it would have acted as a retro-cue, allowing the participants to anticipate which item they should prepare their response for. However, in the new location trials, the cue would point to a previously empty location, making it impossible to prepare their response for any particular item, and therefore making the participants potentially less likely to rely on the cue on those trials. This confusion between Old and New location trials would not have occurred on Cue absent trials, because the participants would have needed to prepare their response for all three of the studied items, irrespective of whether or not the locations had changed. Additionally, this potentially confusing effect of the cue may have only occurred for the cue used in Experiment 4 due to the longer offset between cue onset and probe display onset. In contrast, the cues used in the previous experiments were either presented during the probe display (Experiment 1), or 107ms prior to the probe display onset (Experiment 2 and 3), making it very difficult to gain a retro cue benefit.

One of the more perplexing findings from this set of experiments came from Experiment 5. In Experiment 5, the lack of a Location x Binding interaction indicated that participants were conducting a non-spatial comparison by searching memory for a match with the target item.

However, when a change occurred amongst the uncued distractors, participants were more likely to make a ‘change’ response. The latter effect should not have occurred if the participants were matching only on the basis of the target item, because any mismatch comparisons should have been disregarded. These two findings, from the same experiment, appear to contradict one another. However, previous studies have also found similar effects. For example, Jiang et al. (2000) and Vidal et al. (2005) both found that making a change in an uncued distractor item also increased the proportion of change responses, when compared to a condition where the uncued items did not change. Both of these papers explained this effect by explaining that the comparison process was conducted on the basis of a configuration representation, which cannot be segmented by spatial attention – akin to global in-place matching. However, the findings of Experiment 5 do not fit with this explanation, because if the participants were using global in-place matching, then we should have expected to find a Location x Binding interaction, as in our previous experiments. Therefore, an alternative explanation seems necessary to explain the results in this series of experiments and those of previous studies that have assumed configuration-based matching (e.g., Jiang et al., 2000; Treisman & Zhang, 2006; Vidal et al., 2000).

One possible explanation for these seemingly conflicting results comes from recent evidence that two distinct stages exist during the comparison process (Gilchrist & Cowan, 2014; Hyun et al., 2009; Orzechowski, Necka, & Balas, 2016; Yin et al., 2012). In the first stage, a high capacity filter draws attention to items or locations that have a high likelihood of having changed in the task relevant feature (Hyun et al. 2009; Orzechowski et al. 2016). In the second stage, a limited capacity process performs a slower confirmatory comparison on the candidate item that was selected during the first stage (Hyun et al. 2009; Gilchrist & Cowan, 2014). It is possible that in our Experiment 5, and in the experiments of Jiang et al. (2000) and Vidal et al. (2005), the target change and distractor change competed for attentional selection during the first stage of the comparison process. It is possible that the first stage is relatively automatic and is not influenced by top down information, such as the task relevancy of the items indicated by the cue. As a result, if all else is equal, the distractor and target changes will be selected equally often during the first stage, leading to an inflated change response rate in the second comparison stage. This interpretation does not require the presence of a configural representation, because once the search item is selected in the first stage, a non-spatial search is conducted (Gilchrist & Cowan, 2014), however this account needs to be further tested in future experiments.

One strength of the experiments reported here is that we have replicated effects previously reported in the literature (Treisman & Zhang, 2006). Furthermore, our analysis was supplemented with Bayes factors using the Bayesian ANOVA (Rouder et al., 2012). This type of analysis provides a measure of evidence in support of the null hypothesis, and therefore allowed us to draw conclusions about the null hypotheses, which is not possible using null hypothesis significance testing (e.g., Edwards, Lindman, & Savage, 1963; Rouder et al., 2012). Both types of analyses were largely consistent with one another – significant *p* values were mostly accompanied by Bayes factors favoring the alternative hypothesis, and non-significant *p* values were accompanied by Bayes factors favoring the null. Although there were some examples where interpretation of *p* values and Bayes factors were in contradiction with each other, the effects of interest, such as the Location x Binding interactions, were largely consistent with one another. This consistency between the two types of analyses, and the consistency with which they were replicated across experiments, and previous studies, provides confidence that these effects are real.

How do these results influence how we should think about WWM? The ubiquitous finding within the literature that task-irrelevant locations are used to guide comparisons is often taken as evidence that location is fundamental to the way in which visual features are organized in VWM. Our findings suggest that existing models of VWM should relax their assumption about the extent to which features are necessarily bound to their location. One potentially parsimonious way to incorporate our findings with the existing literature is to assume that observers might automatically encode feature-location binding information (Kondo & Saiki, 2012; Olson & Marshuetz, 2005; Treisman & Zhang, 2006) but that this information does not necessarily always enter the decision-making process. Because location is so highly diagnostic in many everyday tasks, a utilization of location may be the best (or possibly default) choice in many contexts. This view accounts for previous findings, but is also consistent with the view that VWM should be able to flexibly reconfigure the way in which representations are utilized in order to solve novel problems (Oberauer, 2009). Furthermore, it may be the case that individual differences, and environmental factors, will determine what information is accessed during the comparison and decision processes. For example, individual differences in VWM capacity may predict differences in the extent to which irrelevant distractors are erroneously maintained (and thus potentially used in the comparison process) in VWM (Vogel, McCollough & Machizawa, 2005). Additionally, participants who have a better ability to inhibit comparisons with irrelevant items, or to unbind features from their locations, may also be more likely to use a non-spatial comparison strategy. We have identified one possible environmental factor that may modulate the type of comparison strategy which a participant will use - participants may have chosen different strategies, depending on the probability that an irrelevant change had occurred. In the more ‘stable’ environments (Experiments 1, 2, and 4), where irrelevant changes never occurred, participants employed in-place matching, because there was no cost to this strategy. However, in 'noisier' environments, such as when there was a high probability of an irrelevant change (Experiment 5), participants were more likely to use a non-spatial comparison process.

To summarise our findings, a location-based comparison process, such as in-place matching, does not necessarily appear to take place obligatorily, although participants appear to rely on in-place matching by default in the change detection task. Outside of the lab, location may be the best (or default) cue for many every-day tasks, due to its high level of diagnosticity. However, the extent to which location is employed in the comparison process is likely to be determined by the strategic relevance of spatial information in the task. We have shown that when the task demands of the change detection task are changed so as to strongly discourage in-place matching, participants are capable of conducting a non-spatial comparison strategy. In addition, when location is utilized in the comparison process, participants appear to be capable of adapting which items enter this comparison process (Experiment 3). For example, although global in-place matching appears to be the default comparison process, when only some feature-changes are task-relevant, the configuration of the entire display may be used to guide the comparison of only a subset of the items in the display (partial in-place matching). Future developments in conceptualizing VWM therefore should loosen their assumptions about the extent to which location binding occurs automatically or obligatory and take into account how the demands of the task influence the participant's response strategy.

References

Allen, R. J., Castellà, J., Ueno, T., Hitch, G. J., Baddeley, A. D. (2015). What does visual suffix interference tell us about spatial location in working memory? Memory and Cognition, 43, 133-142.

Baddeley, A. D., & Hitch, G. (1974). Working memory. Psychology of learning and motivation, 8, 47-89.

Bayarri, M. J., & García-Donato, G. (2007). Extending conventional priors for testing general hypotheses in linear models. Biometrika, 94, 135-152.

Boduroglu, A., & Shah, P. (2006). Configural representations in spatial working memory. Cognitive Processing, 7, 160-162.

Brainard, D. H. (1997) The Psychophysics Toolbox, Spatial Vision 10, 433-436.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. Psychological review, 83(1), 1.

Bundesen, C. (1990). A theory of visual attention. Psychological review, 97(4), 523.

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. Proceedings of the National Academy of Sciences, 98, 5363-5367.

Chen, Z. (2009). Not all features are created equal: Processing asymmetries between location and object features. Vision research, 49,1481-1491.

Currie, C. B., McConkie, G. W., Carlson-Radvansky, L. A., & Irwin, D. E. (2000). The role of the saccade target object in the perception of a visually stable world. Perception & Psychophysics, 62, 673-683.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. Annual review of neuroscience, 18, 193-222.

Donkin, C., & Nosofsky, R. M. (2012). The structure of short-term memory scanning: An investigation using response time distribution models. Psychonomic Bulletin & Review, 19, 363-394.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. Psychological Review, 70, 193-242.

Eimer, M., & Grubert, A. (2014). Spatial attention can be allocated rapidly and in parallel to new visual objects. Current Biology, 24, 193-198.

Gilchrist, A. L., & Cowan, N. (2014). A two-stage search of visual working memory: investigating speed in the change-detection paradigm. Attention, Perception, & Psychophysics, 76, 2031-2050.

Golomb, J. D., Kupitz, C. N., & Thiemann, C. T. (2014). The influence of object location on identity: A “spatial congruency bias”. Journal of Experimental Psychology: General, 143, 2262.

Guérard, K., Morey, C. C., Lagacé, S., & Tremblay, S. (2013). Asymmetric binding in serial memory for verbal and spatial information. Memory and Cognition, 41, 378-391.

Hayes, S. M., Nadel, L., & Ryan, L. (2007). The effect of scene context on episodic object recognition: parahippocampal cortex mediates memory encoding and retrieval success. Hippocampus, 17, 873-889.

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. Cognitive Psychology, 10, 302-323.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32, 58-69.

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. Journal of Experimental Psychology: Human Perception and Performance, 33, 31-47.

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. Visual Cognition, 17, 273-291.

Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. Cognition, 113, 150-166.

Hollingworth, A., & Rasmussen, I. P. (2010). Binding Objects to Locations: The Relationship between Object Files and Visual Working Memory. Journal of Experimental Psychology: Human Perception and Performance, 36, 543-564.

Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. Journal of Experimental Psychology: Human Perception and Performance, 35, 1140.

Jeffreys, H. (1961). Theory of Probability (3rd Ed.). Oxford Classic Texts in the Physical Sciences. Oxford, UK: Oxford University Press.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). The Organisation of Visual Short-Term Memory. Journal of Experimental Psychology: Learning, Memory and Cognition, 26, 683-702.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye. Attention and performance, 9, 187-203.

Kondo, A., & Saiki, J. (2012). Feature-specific encoding flexibility in visual working memory. PloS one, 7.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390, 279-281.

Mou, W., Xiao, C., & McNamara, T. P. (2008). Reference directions and reference objects in spatial memory of a briefly viewed layout. Cognition, 108, 136-154.

Oberauer, K. (2009). Design for a working memory. Psychology of learning and motivation, 51, 45-100.

Olivers, C., & Schreij, D. (2011). Object representations maintain attentional control settings across space and time. Journal of Vision, 11, 141-141.

Olson, I. R., & Marshuetz, C. (2005). Remembering “What” brings along “Where” in visual working memory. Perception & Psychophysics, 67, 185-194.

Olson, I. R., Moore, K. S., & Drowos, D. B. (2008). The contents of visual memory are only partly under volitional control. Memory & cognition, 36(7), 1360-9.

Orzechowski, J., Nęcka, E., & Balas, R. (2016). Task conditions and short-term memory search: two-phase model of STM search. Polish Psychological Bulletin, 47, 12-20.

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. Journal of Experimental Psychology: Human Perception and Performance, 16, 332.

Papenmeier, F., Huff, M., & Schwan, S. (2012). Representation of dynamic spatial configurations in visual short-term memory. Attention, Perception, & Psychophysics, 74, 397-415.

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, Spatial Vision 10, 437-442.

Pertzov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. Attention, Perception, & Psychophysics, 76, 1914-1924.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rajsic, J., & Wilson, D. E. (2014). Asymmetrical access to color and location in visual working memory. Attention, Perception, & Psychophysics, 76, 1902-1913.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. Journal of Mathematical Psychology, 56, 356-374.

Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. Cognitive Psychology, 14, 353-409.

Silvis, J. D., & Shapiro, K. L. (2014). Spatiotemporal configuration of memory arrays as a component of VWM representations. Visual Cognition, 22, 948-962.

Simonsohn, U. (2015). Small Telescopes Detectability and the Evaluation of Replication Results. Psychological science, 0956797614567341.VA designs. Journal of Mathematical Psychology, 56, 356-374.

Simonsohn, U. (2015). Small telescopes detectability and the evaluation of replication results. Psychological Science.

Sternberg, S. (1966). High-speed scanning in human memory. Science, 153, 652-654.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. American scientist, 57, 421-457.

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. Memory & Cognition, 34, 1704-1719.

Van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. Psychological review, 121, 124.

Vicente-Grabovetsky, A., Carlin, J. D., & Cusack, R. (2012). Strength of retinotopic representation of visual memories is modulated by strategy. Cerebral Cortex, 313.

Vidal, J. R., Gauchou, H. L., Tallon-Baudry, C., & O'regan, J. K. (2005). Relational information in visual short-term memory: The structural gist. Journal of Vision, 5, 8.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. Nature, 438(7067), 500-3.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. Journal of Experimental Psychology: Human Perception and Performance, 27(1), 92-114.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in Short-Term Visual Memory. Journal of Experimental Psychology: General, 131(1), 48-64. Elsevier.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. Journal of vision, 4, 11-11.

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2012). Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. Visual Cognition, 20, 1-28.

Yin, J., Gao, Z., Jin, X., Ding, X., Liang, J., & Shen, M. (2012). The neural mechanisms of percept–memory comparison in visual working memory. Biological psychology, 90, 71-79.

Zellner, A., & Siow, A. (1980). Posterior odds ratios for selected regression hypotheses. Trabajos de estadística y de investigación operativa, 31, 585-603.

Figures

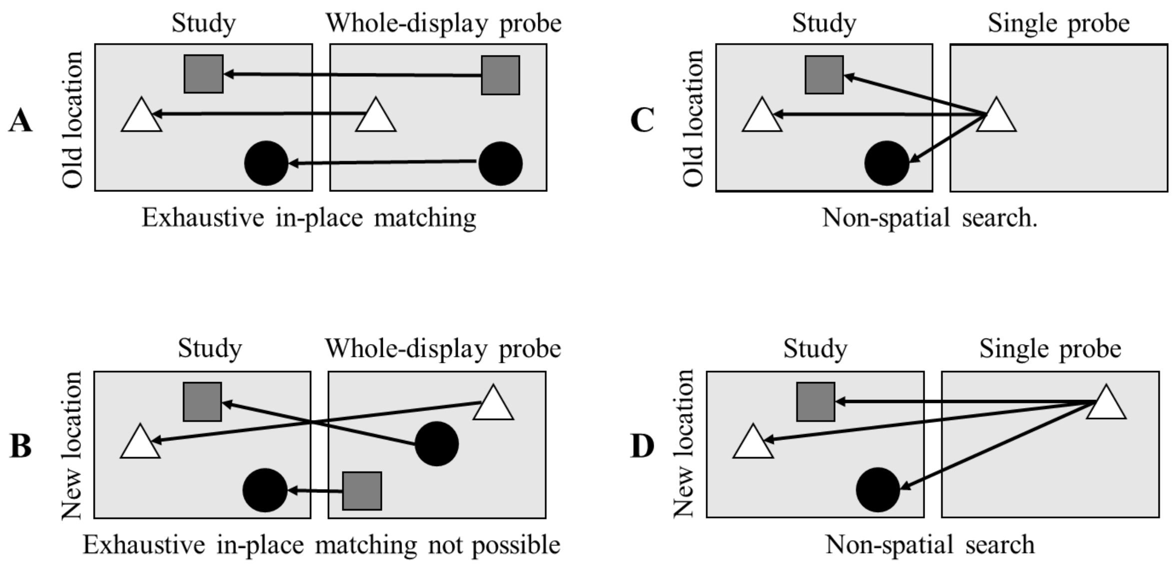


Figure 1. Examples of possible comparison processes under different types of probe display. **A.** Whole-display probe with items maintaining their original location; participants can use exhaustive in-place matching. Responses are based on the detection of a miss-match comparison. **B.** Whole-display probe in new locations; participants are unable to use in-place matching, and must employ some other comparison process. **C.** Single probe in original location; participants may exhaustively search memory, with responses based on match detection. **D.** Single probe in a new location; participants cannot make use of location, and therefore must serially search memory, making comparisons with items at multiple locations. Responses are based on the detection of a match comparison.

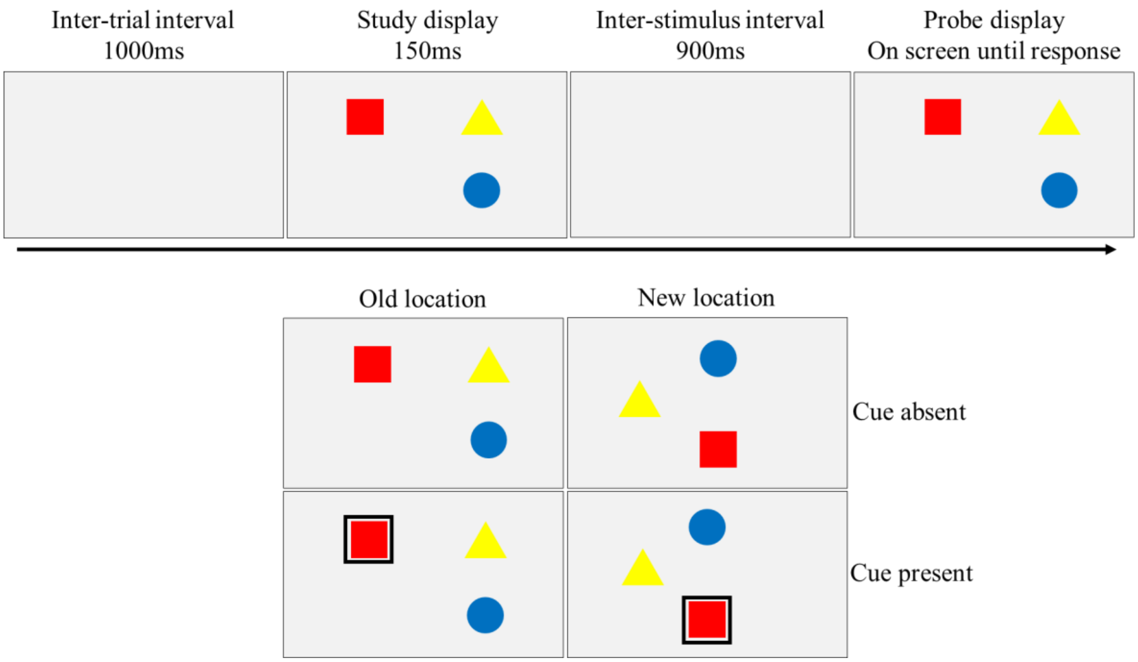


Figure 2. A schematic of an example trial, with four possible probe displays. Examples of Switched binding, or New feature trials are not depicted.

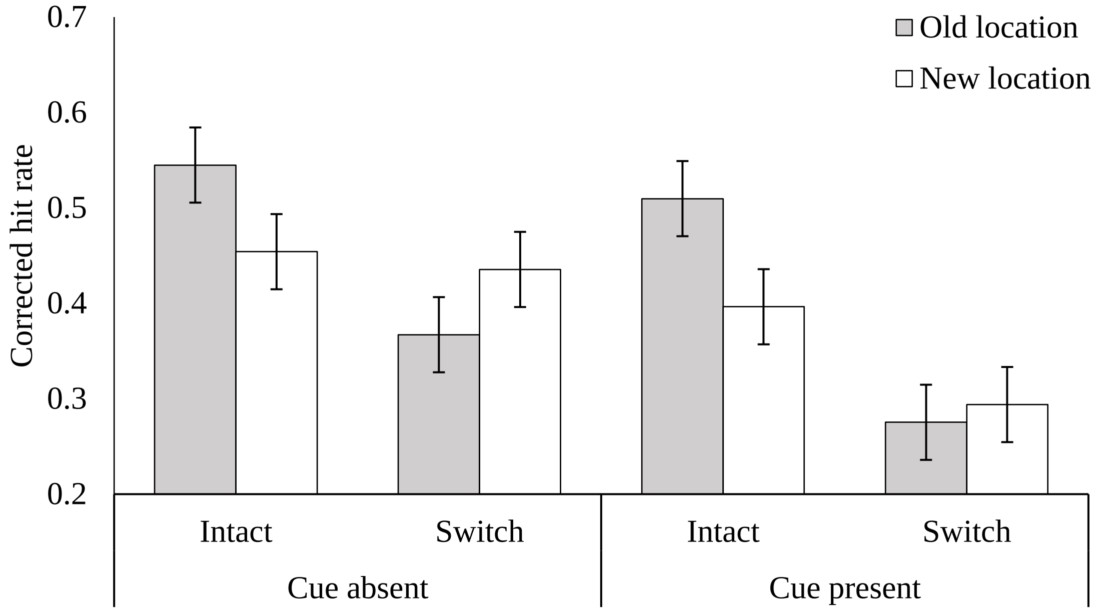


Figure 3. Corrected hit rates for the Location by Binding by Cue type interaction in Experiment 1.

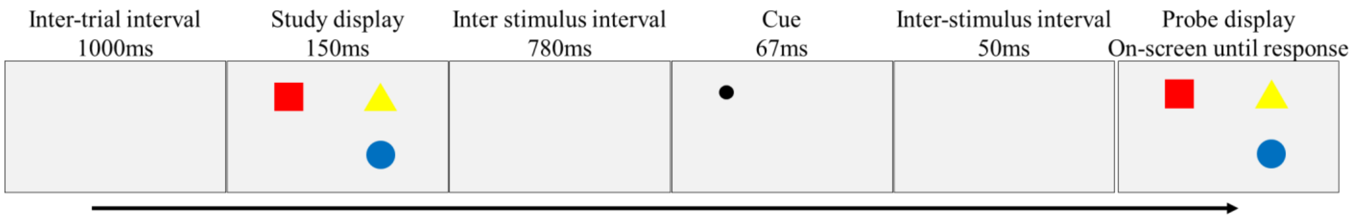


Figure 4. A schematic of a typical trial in Experiment 2. This trial shows an example of a Cue present, Match feature, Intact binding, Old location trial.

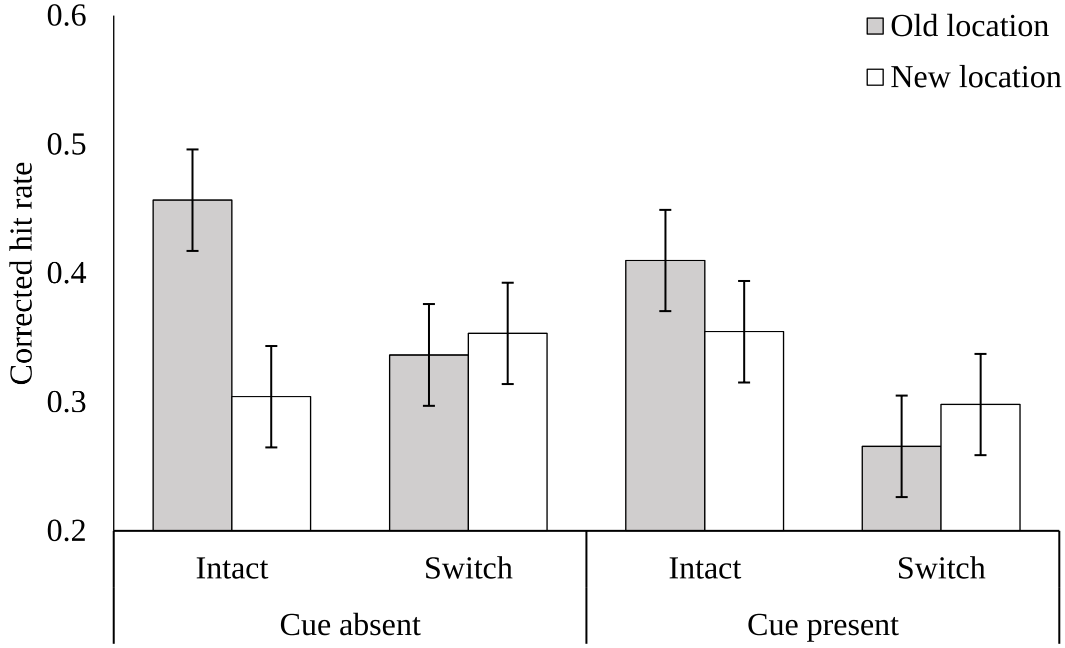


Figure 5. Corrected hit rates in each of the Cue, Location, and Binding conditions, in Experiment 2.

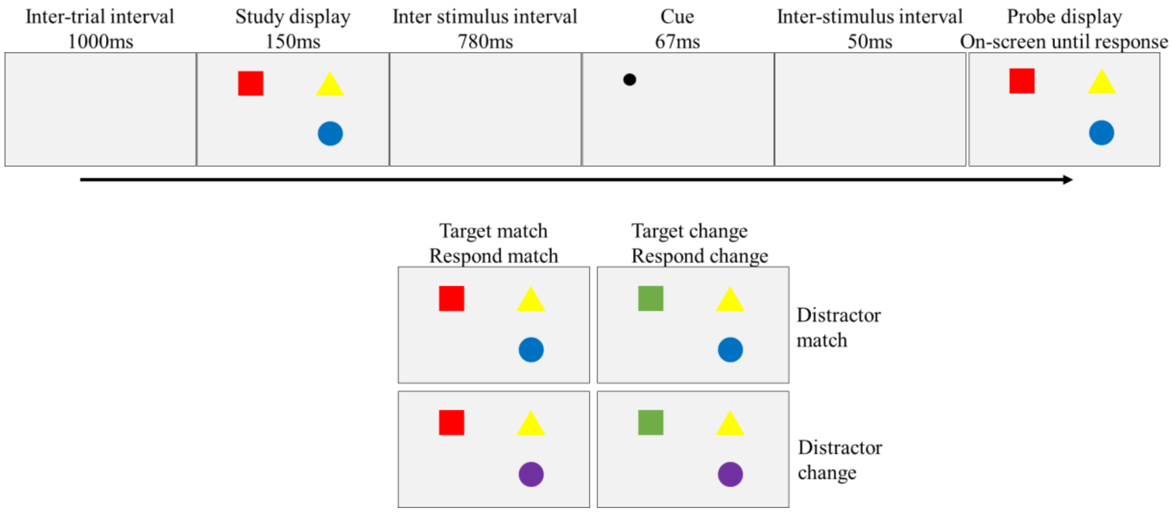


Figure 6. Top: Schematic of an example trial in Experiment 3. Bottom: Examples of probe displays with combinations of Target match/change and Distractor match/change, along with the correct response.

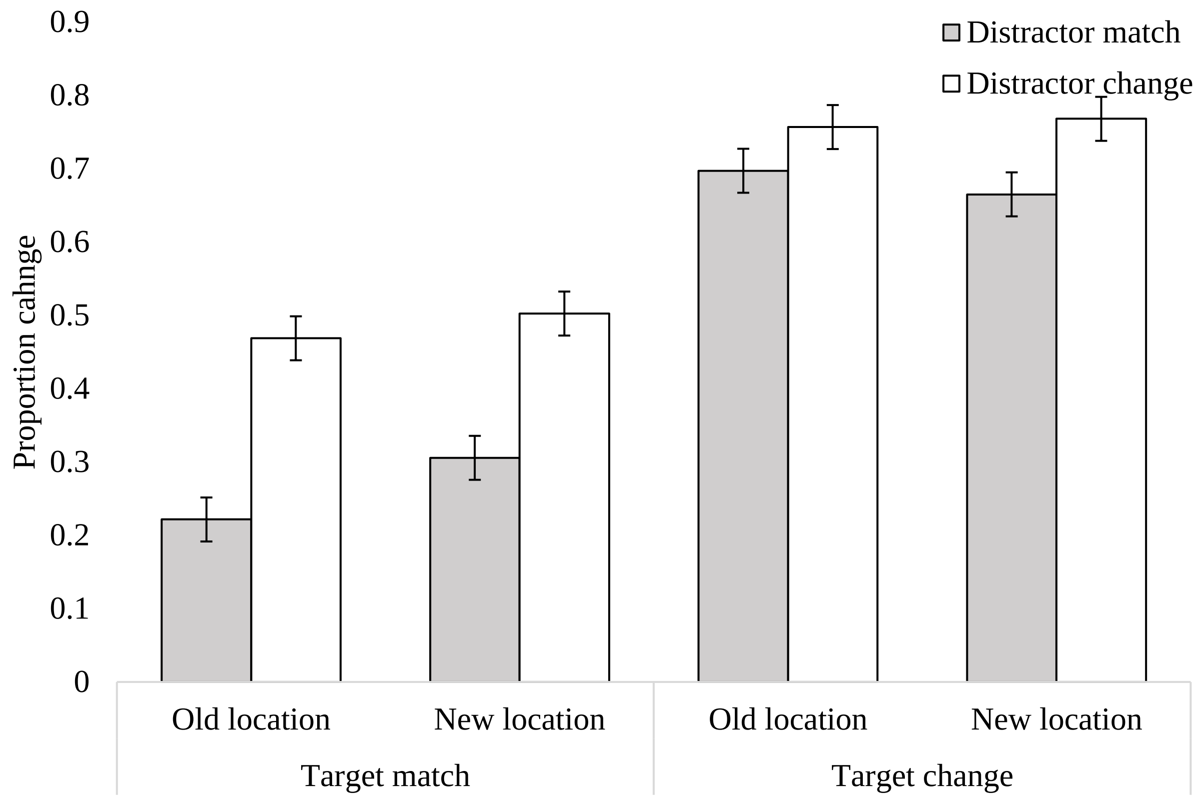


Figure 7. Mean proportion change for each condition in Experiment 3.

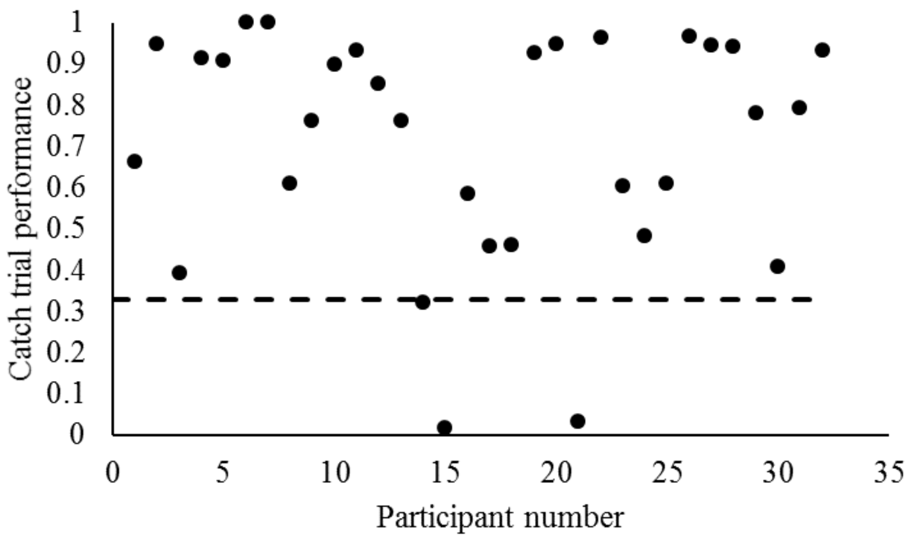


Figure 8. Each point represents the mean proportion of correct responses on catch trials for that participant.

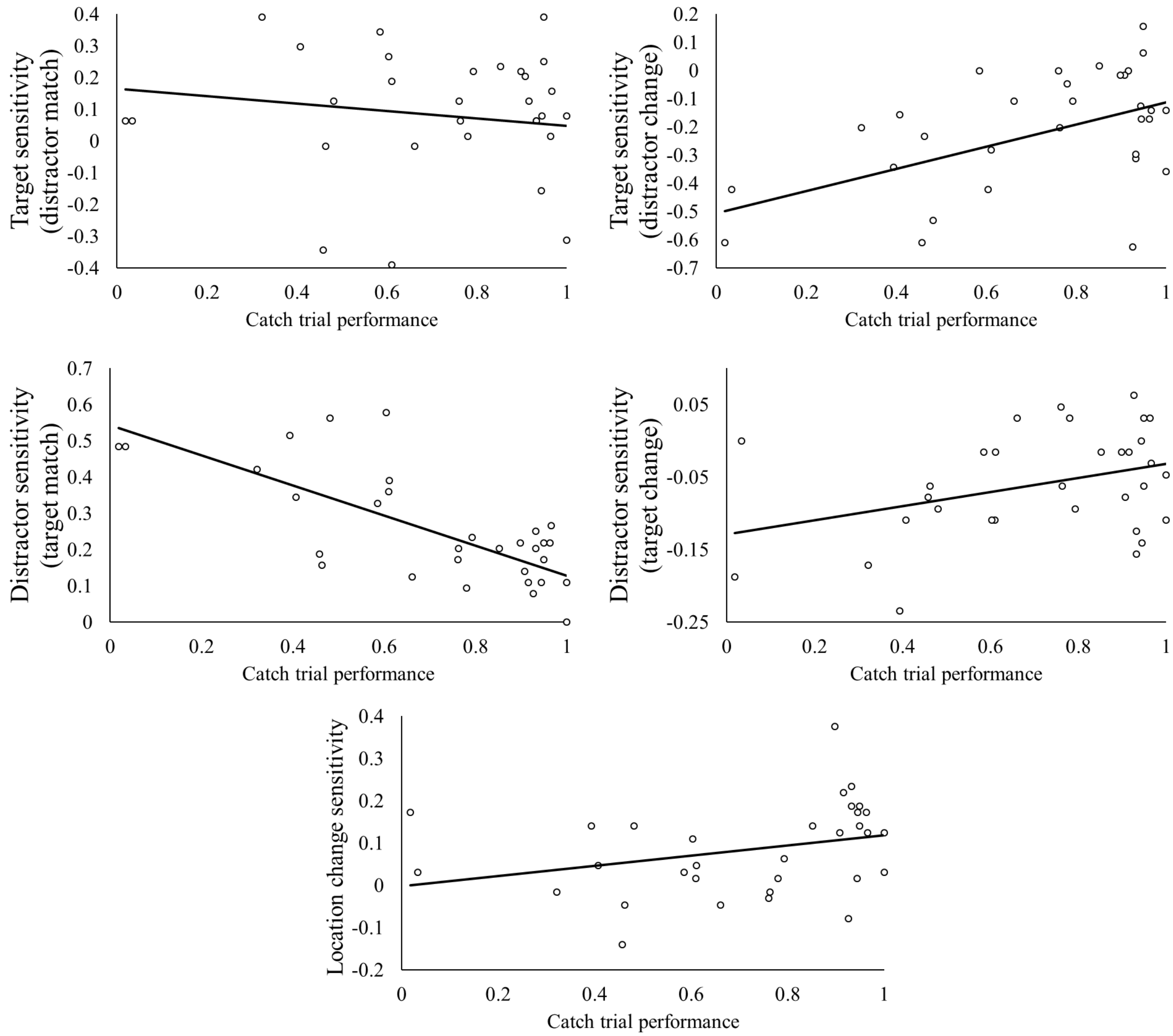


Figure 9. Correlations between mean performance for individual participants on the catch trials, and mean their mean sensitivity to different conditions in the change-detection task.

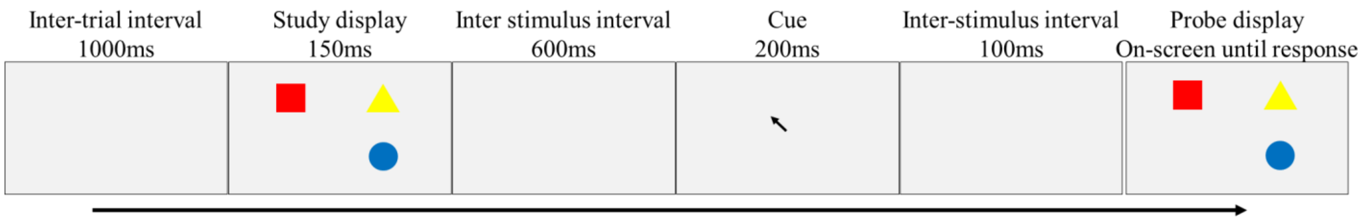


Figure 10. A schematic of a typical trial in Experiment 4. This trial shows an example of a Cue present, Match feature, Intact binding, and Old location trial.

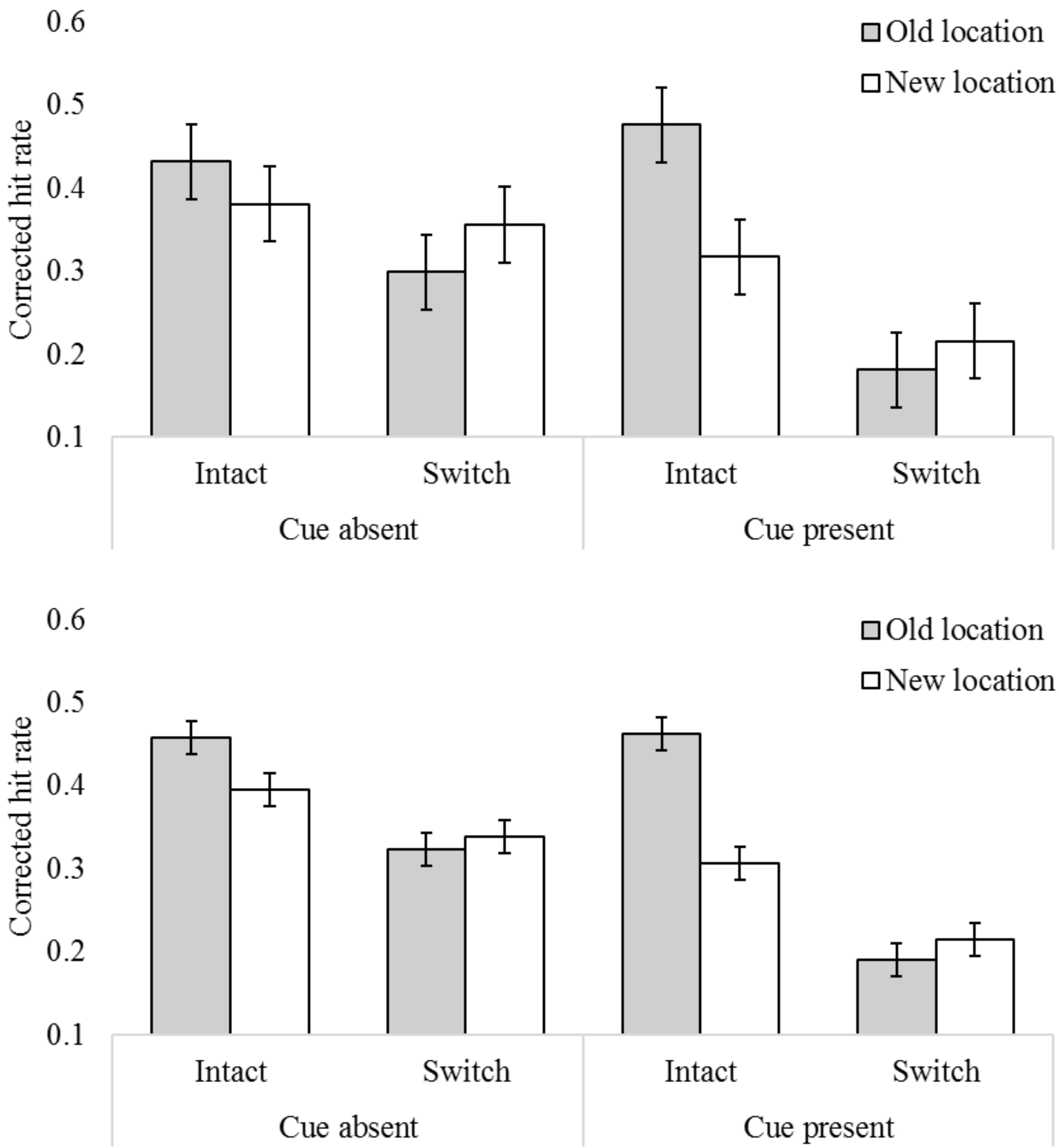


Figure 11. Corrected hit rates in each of the Cue, Location, and Binding conditions in Experiment 4. Top: original study. Bottom: Replication.

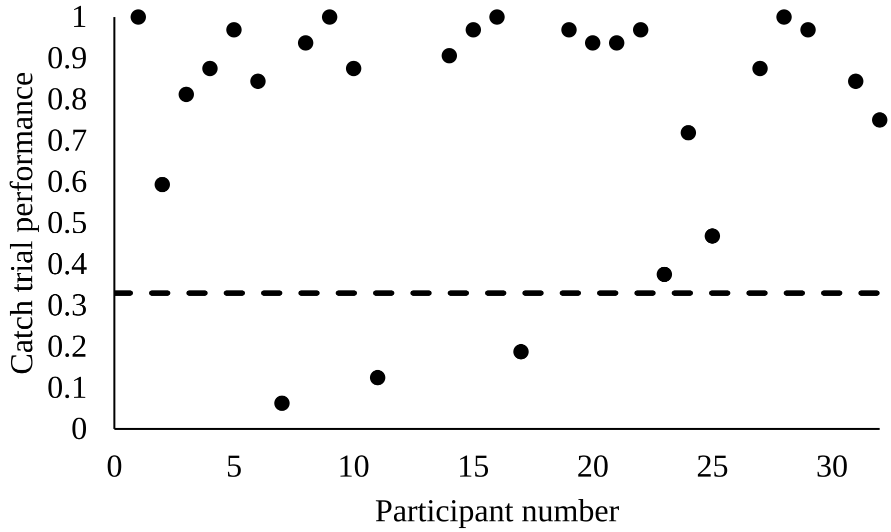


Figure 12. Average catch-trial performance for each participant.

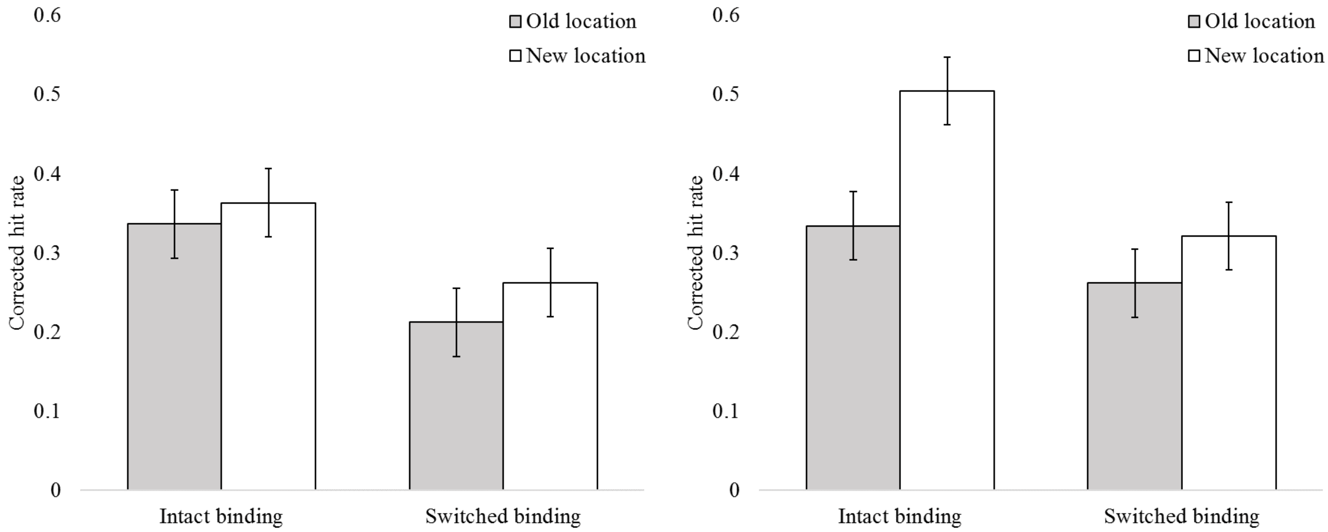
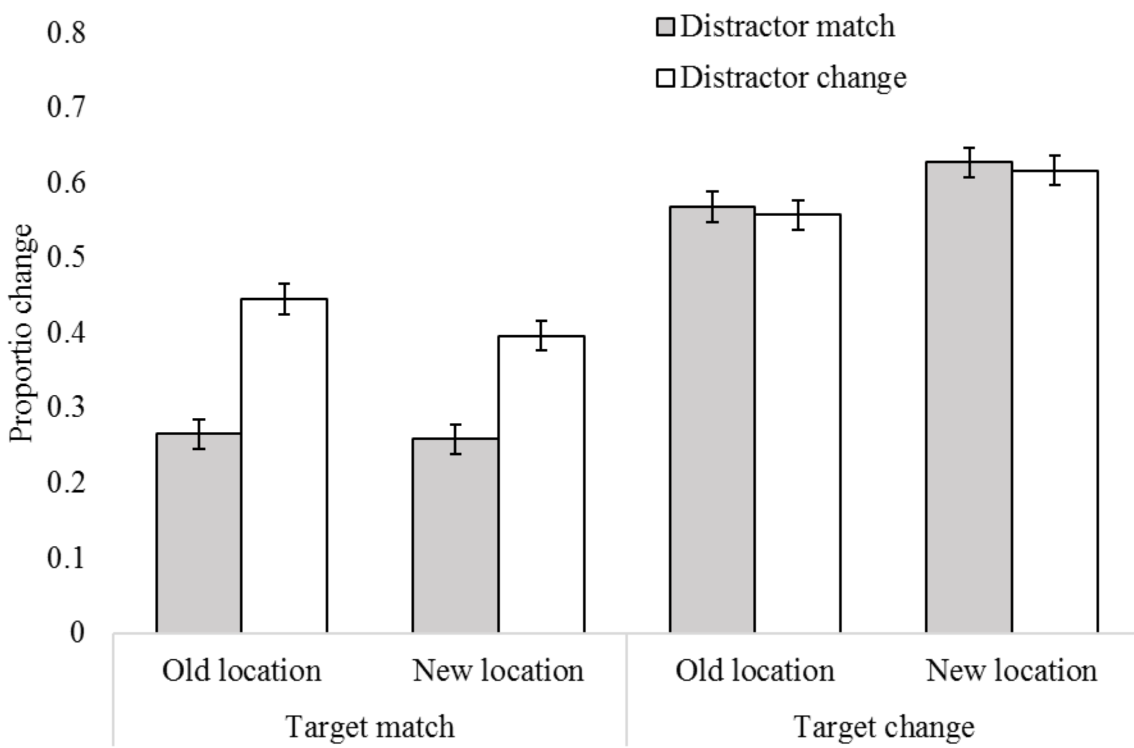


Figure 13. Mean corrected hit rate in each of the Binding and Location conditions in Experiment 5. Left: Original experiment. Right: Replication experiment.



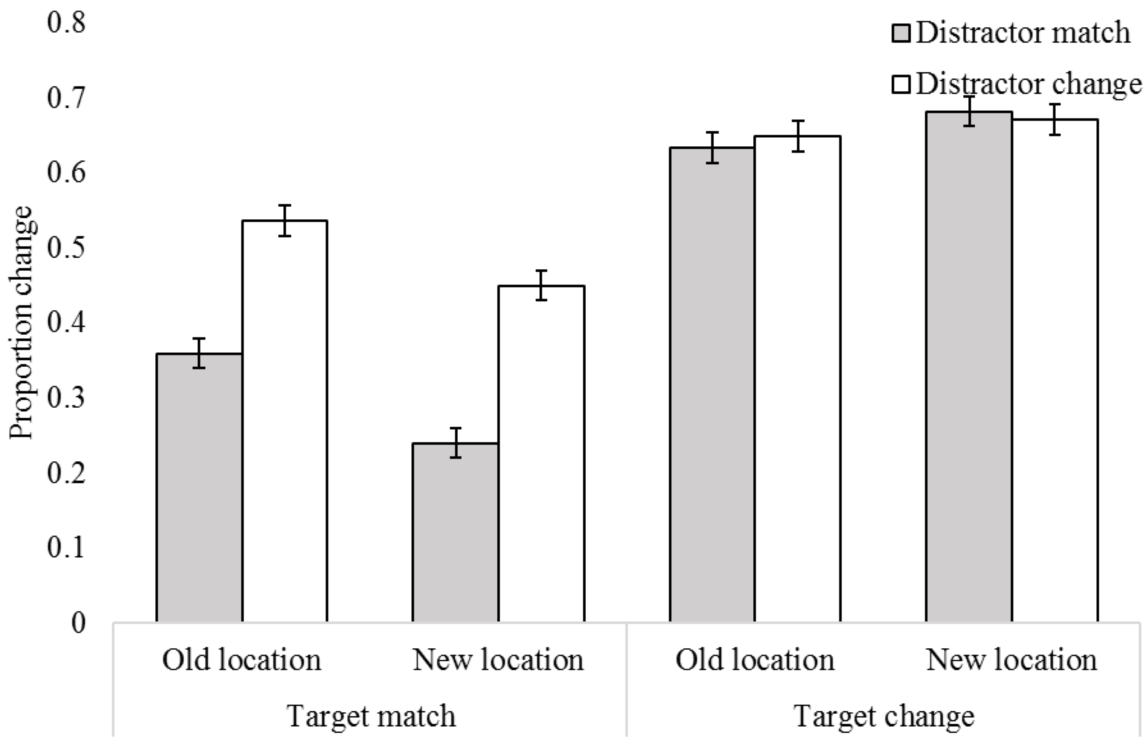


Figure 14. Mean proportion change for each condition in Experiment 5. Top: Original experiment. Bottom: Replication experiment.

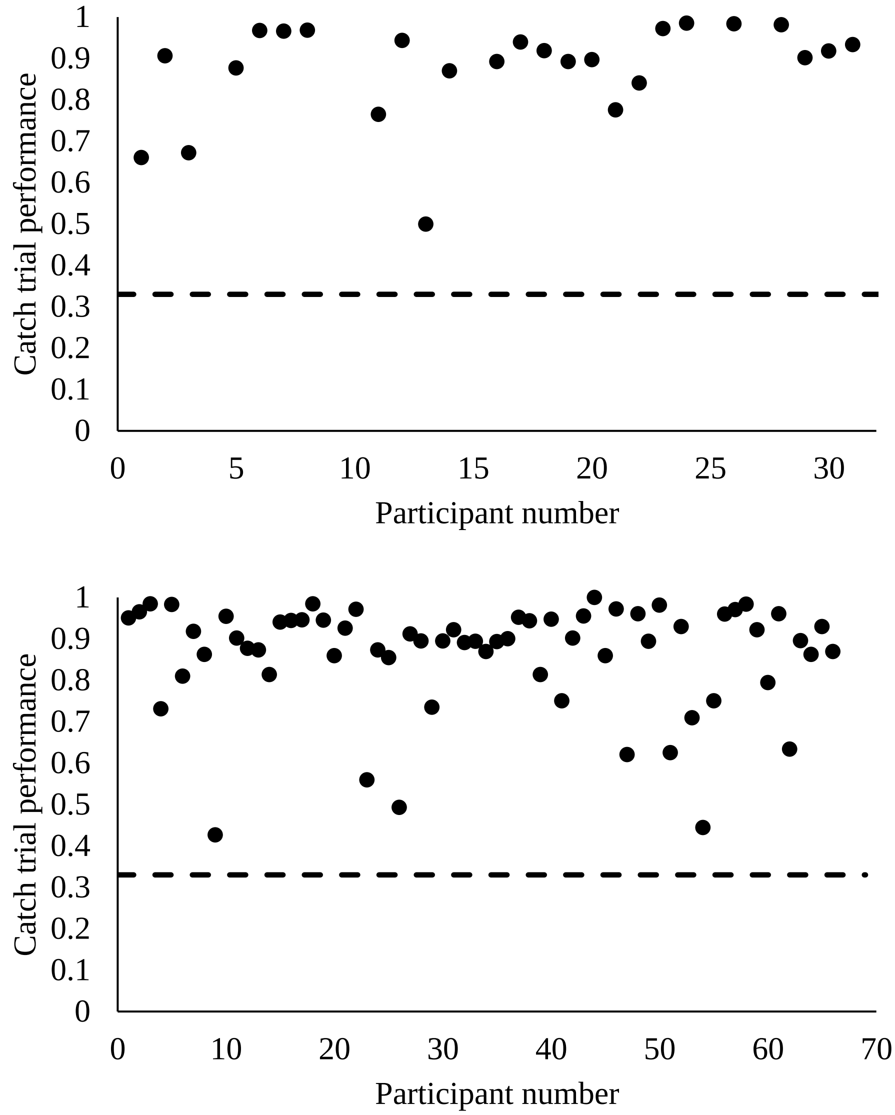


Figure 15 shows the average catch-trial performance for each participant. The dashed line represents chance performance – the probability of selecting the target location, if they randomly selected one of the three possible probe item locations. However, it was possible for participants to select any location on the screen, and empty screen locations were selected on 3.68% of trials.